Volume 58, Parts 1 and 2 *May 1967*

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
ARTHUR SUMMERFIELD

Br. J. Psychol. (1967), **58**, 1 and 2

Published January 1968

We and the Editor regret the delay in the publication of this issue.

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, P.O. BOX 92, 200 EUSTON ROAD, LONDON, N.W. 1

AMERICAN BRANCH:

32 EAST 57TH STREET, NEW YORK, N.Y. 10022

60s. net $*11.00*

Assistant Editors

JAMES HOGG PETER MITTLER

Preparation of Papers for the British Journal of Psychology

In submitting papers for publication in the *British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. **56** (1965) onwards.

A brief summary, suitable for use as an abstract, should precede the article.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26** 231–48.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *L'Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

For other general guidance, authors are referred to *The Journal of Physiology: Suggestions to Authors.* Cambridge University Press. Price 2*s*.

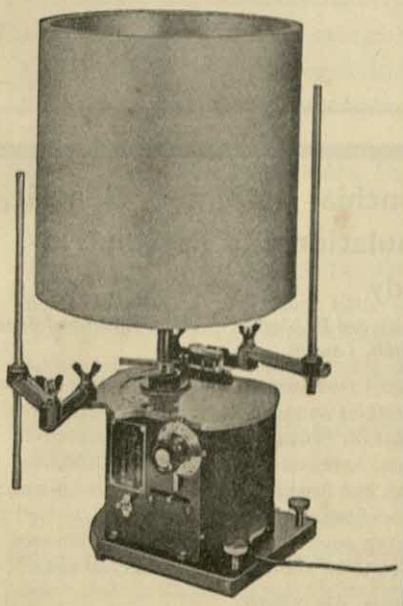
The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

PALMER

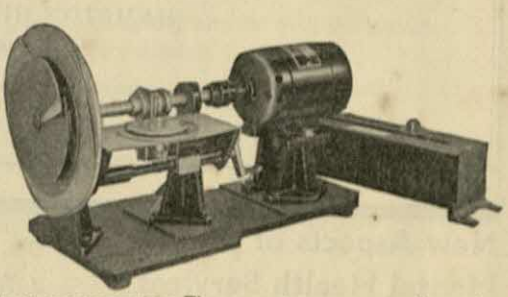
32

A full range of equipment for psychology laboratories

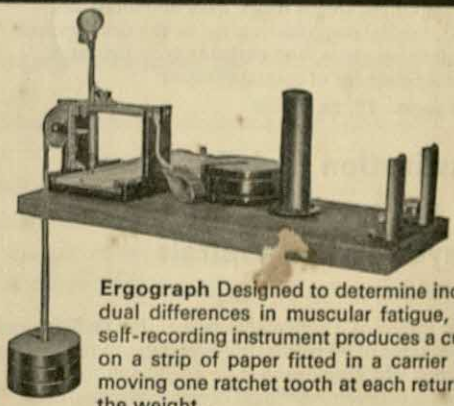
C. F. Palmer (London) Ltd. can supply instruments and accessories for measuring and testing muscle movement, muscular fatigue, motor fatigue, auditory capacity, reaction to colour, response times and other variables measured in psychology laboratories.
Three instruments from the extensive range of C. F. Palmer products are illustrated.
For further details and a copy of the catalogue write to C. F. Palmer (London) Ltd. at the address below.



Electric Twelve Kymograph Enables movements of various kinds to be represented accurately on a permanent paper record. A 12-speed gear-box gives variable control between ·8 seconds and 13 hours a revolution.



Colour comparator The operator can vary the colour proportions of the smaller disc from 10° to 350° while it is revolving, and compare it with the larger disc, which revolves at the same speed.



Ergograph Designed to determine individual differences in muscular fatigue, this self-recording instrument produces a curve on a strip of paper fitted in a carrier tray moving one ratchet tooth at each return of the weight.

CFP

C. F. Palmer (London) Ltd., Myographic Works, Effra Rd., Brixton, London, S.W.2
Telephone: 01-733 2173/5
A member of the Baird & Tatlock Group of Companies
CP.2

ELECTRONIC DEVELOPMENTS

Apparatus for Psychological Research

TACHISTOSCOPES

PSYCHOGALVANOMETERS PURSUIT ROTORS

VOICE KEYS TIMERS PULSE GENERATORS

LAMP CONTROLLERS ILLUSION FRAMES

MAGNETIC RECORDING EQUIPMENT

Send for list to:

ELECTRONIC DEVELOPMENTS
46 ST JAMES'S AVENUE
HAMPTON HILL, MIDDLESEX
(01-979 5047)

New Aspects of the Mental Health Services

H. Freeman, *Salford Royal Hospital,* **and J. Farndale,** *City of Westminster College, London*

A sequel to the authors' *Trends in the Mental Health Services,* this book is a collection of papers reviewing the important trends which have developed since 1963. Special emphasis is given to the developments in other countries, and particularly to the extra-mural treatment of mental illnesses.

640 pages. £5. 5s./$15.00

Institution and Outcome: A Comparative Study of Psychiatric Hospitals

L. P. Ullman, *University of Illinois*

The results of a systematic study of the facilities and operations of 30 psychiatric hospitals are reported in a book which permits a close analysis of the association between individual hospital characteristics and their effectiveness. The conclusions drawn from this study having been viewed from both the organizational and behavioural aspects lead to useful suggestions for improvements.

216 pages. 50s./$7.50

Bronchial Asthma-A Genetic, Population and Psychiatric Study

D. Leigh and E. Marley, *Bethlem Royal and Maudsley Hospitals, London*

In order to investigate the various aspects of bronchial asthma an extensive survey was carried out amongst the patients of two general practices in London. Approximately 1,700 individuals were studied, and from their records the heredity and prevalence of asthma has been determined, as well as numerous points concerning a range of illnesses believed to be associated with bronchial asthma.

208 pages. 63s./$10.00

PERGAMON PRESS

Headington Hill Hall, Oxford.
44-01 21st Street, Long Island City,
New York 11101, U.S.A.

ZEITSCHRIFT FÜR EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

Herausgeber für Deutschland
Prof. Dr. PH. LERSCH, München; Prof. Dr. F. SANDER, Bonn
Prof. Dr. K. GOTTSCHALDT, Göttingen

Herausgeber für Österreich
Prof. Dr. H. ROHRACHER, Wien

Herausgeber für die Schweiz
Prof. Dr. R. MEILI, Bern

Redaktion, Dr. C. J. HOGREFE, Göttingen

BAND XIII-3

Wolfram Ehlers und Rudolf Cohen. Eine Untersuchung zur Adaption an die verzögerte akustische Rückmeldung der Sprache unter verschiedenen Instruktionen.

Heinz Franke. Bemerkungen zur Problematik der sogenannten unterschwelligen Wahrnehmung.

Klaus Holzkamp und Erich Perlwitz. Absolute oder relative Größenakzentuierung?

Ludwig Kötter und Angelika Keilig. Zum Einfluß geistiger Belastung und Belastungserwartung auf die Flimmerverschmelzungsfrequenz bei Universitätsstudenten.

Stanley Milgram. Einige Bedingungen von Autoritätsgehorsam und seiner Verweigerung.

Mervyn William Perrine. Bezugsskalenbildung, Diskrimination und verbale Verstärkung.

Franz Schönberger. Soziometrische Untersuchungen über zerebral gelähmte Kinder in einer Körperbehindertenschule.

Herbert Selg und Gottfried Lischke. Eine Faktorenanalyse von Aggressionsvariablen.

BAND XIII-4

Friedrich Denig. Über das Erlerhen einer Reiz-Reaktions-Verbindung bei ähnlichen Reizen.

Theo Herrmann und Kurt Stapf. Untersuchungen zum Erlernen von Wahlreaktionen. Teil II: Zur Theorie leistungsrelevanten Lernverhaltens.

Eberhard Kunkel. Die Einschätzung der relativen Gefährlichkeit verschiedener Verkehrsdelikte.

Fritz Reimer und Kurt Hauss. Provozierte Trugwahrnehmungen im psychologischen Experiment.

Günther Reinert, Paul B. Baltes und Lothar R. Schmidt. Kritik einer Kritik der Differenzierungshypothese der Intelligenz.

Lutz v. Rosenstiel. Zur Frage der Angst- und Feindseligkeitsinhalte in Formdeuteverfahren.

Klaus Schüttler-Janikulla und Heinz Krohne. Transferprobleme und Schulreifetraining.

Johannes Franz Ullmann und Ingeborg Maria Ullmann. Farbspezifische Helligkeit beim direkten und indirekten, beim Tages- und Dämmerungssehen.

All articles are provided with summaries in English and French.
The journal appears quarterly and contains 160 pages per number.
Subscription is DM 64.— per volume.

VERLAG FÜR PSYCHOLOGIE . DR. C. J. HOGREFE . GÖTTINGEN
(Brentanoweg 10, Germany)

OCCUPATIONAL PSYCHOLOGY

Editor: ALEC RODGER

OCTOBER 1966 Volume 40, No. 4

CONTENTS

Psychology and Military Proficiency. By N. A. B. WILSON.

Concepts of Work. By SYLVIA SHIMMIN.

Some Characteristics of Foremen in the Iron and Steel Industry. By PETER B. WARR, MICHAEL BIRD and RAYMOND W. HADFIELD.

Frequent Alternation of Shifts on Continuous Work. By J. WALKER.

Age and Internal Labour Mobility of Semi-Skilled Workers. By C. K. ELLIOTT.

Age, Speed and Accuracy: A Study in Industrial Re-Training. By G. H. JAMIESON.

JANUARY 1967 Volume 41, No. 1

CONTENTS

'Technical' Change and the Misuse of Professional Manpower: Some Studies and Observations. By DENIS PYM.

Why Ergonomics? By K. F. H. MURRELL.

Comments on Murrell's 'Why Ergonomics?' By W. T. SINGLETON, B. SHACKEL, K. F. JACKSON and E. N. CORLETT.

Extending the Occupational Environment: The Measurement of Organisations. By KERR INKSON, ROY PAYNE and D. PUGH.

The Evaluation of Visual Displays: Techniques and Trends. By J. M. ROLFE.

Isolation and Description of Occupational Stereotypes. By JOAN GUILFORD.

BOOK REVIEWS.

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology

14 Welbeck Street, London, W.1

The British Journal of CRIMINOLOGY

DELINQUENCY AND DEVIANT SOCIAL BEHAVIOUR

Formerly THE BRITISH JOURNAL OF DELINQUENCY

Editors:

EDWARD GLOVER · HERMANN MANNHEIM · EMANUEL MILLER

EDITORIAL ARTICLES

CURRENT SURVEY

RESEARCH AND METHODOLOGY CURRENT NOTES

BOOKS AND PERIODICALS

Published on behalf of

THE INSTITUTE FOR THE STUDY AND TREATMENT OF DELINQUENCY

Annual Subscription £3. 17s. 6d.

Quarterly 21s. net

11 NEW FETTER LANE · LONDON · E.C.4

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Edited by A. J. WATSON

VOL. XIX, Part 1 February 1967 20*s. net*

A. TREISMAN and G. GEFFEN. Selective attention: perception or response?

J. STEINER. Observing responses and uncertainty reduction.

R. DAVIS and D. H. TAYLOR. Classification on the basis of conditional cues.

P. M. A. RABBITT and S. PHILLIPS. Error-detection and correct latencies as a function of S-R compatibility.

A. W. PRESSEY. The distance gradient in the after-effect of kinaesthetic width judgement.

J. K. CLARKSON. Proactive inhibition as an artefact of the method of paced anticipation.

Shorter Articles and Notes

D. R. HASLAM and E. A. C. THOMAS. An optimum interval in the assessment of pain threshold.

D. SHANKWEILER and M. STUDDERT-KENNEDY. Identification of consonants and vowels presented to left and right ears.

D. J. R. LAMING. On procuring human subjects.

G. S. BRINDLEY, R. H. S. CARPENTER and D. N. RUSHTON. Reaction times for simple shape discriminations requiring one or both visual cortices.

J. A. NORTH, D. A. GRANT and R. A. FLEMING. Choice reaction time to single digits, spelled numbers and 'right' and 'wrong' arithmetic problems and short sentences.

Apparatus

A. W. F. HUGGINS. Accurate delays for auditory feedback experiments.

Proceedings of the Experimental Psychology Society, 1966.

Book Reviews.

Publications Received.

Published for the Experimental Psychology Society by

W. HEFFER & SONS LTD, CAMBRIDGE

Annual Subscription, 60s. to W. Heffer & Sons Ltd, or any bookseller

REVUE DE PSYCHOLOGIE APPLIQUÉE

PUBLICATION TRIMESTRIELLE

Directeurs: Dr P. PICHOT et P. RENNES

4e trimestre 1966, Volume 16, No. 4

SOMMAIRE

F. CLEMENT. — Analyse de l'Échelle de Mémoire de Wechsler. Facteurs qui influent sur ses résultats.

S. HUGOT. — Essai d'adaptation du Test du Village au milieu sénégalais.

S. PACAUD. — Portrait. Notice biographique et bibliographique.

1e trimestre 1967, Volume 17, No. 1

SOMMAIRE

P. PICHOT, P. NICOLAS-CHARLES et J. LION. — Étude factorielle de quatre tests psycho-moteurs.

C. DE CAE. — Comparaison expérimentale de deux méthodes d'étalonnage à deux dimensions.

M. DE MOURA SIMOES. — Résultats d'adolescents épileptiques au Test de Wechsler-Bellevue.

B. F. SKINNER. — Portrait, Notice biographique et bibliographique.

Rédaction et Administration: Square Jouvenet, PARIS, 16°

ABONNEMENTS: (un an) France: 15 F; Étranger: 18 F

le numéro: France: 4 F; Étranger: 5 F

Règlements au nom de la Sté André VIDAL & ASSOCIÉS

15, rue Henri-Heine, PARIS 16°, C.C.P. PARIS 20 272 53

NUMÉRO SPÉCIMEN SUR DEMANDE

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA

YEAR XXVII. ISSUE IV–V JULY–OCTOBER 1966

CONTENTS

LEADING ARTICLE

PAPERS

L. ANCONA — I. SANGUINETI. Psychosomatic medicine.

G. CANEPA. Contribution of personality theories to criminological research. Clinical features.

M. BLEULER. The meaning of psychotherapeutic research with regard to theory about schizophrenia and to the schizophrenic patient.

H. A. WITKIN. Some implications of research on cognitive style for problems of education.

L. RIGO. Diagnostical r.é.d. (rêve éveillé dirigé) and Rorschach.

G. GIOVANELLI — P. TAMPIERI — A. VENTURINI. Experimental inquiry about relations between anxiety and learning.

V. FAENZA. Equivocity conditions of answer with regard to the problem of 'subliminar' perception.

G. ALBERTI. New problems regarding chronic patients in present assistance organizations.

C. FANELLI — C. DAMONTE — G. CLERICI. Preliminary observations about juvenile problems in the fields of affective dynamics and ethical behaviour.

NOTES AND DISCUSSIONS

A. SEMERARI — A. CASTELLANI. Alterity and alienity.

YEAR XXVII. ISSUE VI NOVEMBER–DECEMBER 1966

CONTENTS

LEADING ARTICLE

PAPERS

L. ANCONA. 'Toning' between autoplastic dynamics and alloplastic dynamics.

G. PERI. Inquiry about a few features of interpersonal valuation.

E. DE RENZI — P. FAGLIONI. Influence of sleep privation and working upon yield in watching tests.

M. BACCIAGALUPPI. Social psychiatry: a critical review.

Every article is followed by an Italian, French, English and German summary.
Every issue summarizes many Italian and Foreign works of psychology, neurology and psychiatrics.
Annual subscription: Italy, 3,000 Lire; foreign countries, 6,000 Lire (or equivalent foreign money).
Redaction — Administration: 1, Largo Gemelli, Milan (Italy).

Bureau of Ednl. & Psyl. Research
(S. C. E. R. T.)
Date .....
Acc. J. 897

BRAIN (A Journal of Neurology)

Published Quarterly 20s. net. Annual Subscription £4. 4s. Post free

Contents of the latest issue:

Disorders of Memory. By SIR CHARLES SYMONDS.

Infarction of the Spinal Cord. By HUGH GARLAND, J. GREENBERG and D. G. F. HARRIMAN.

Vital Staining and Electron Microscopy of the Intramuscular Nerve Endings in the Neuropathy of Adult Cœliac Disease. By W. T. COOKE, A. G. JOHNSON and A. L. WOOLF.

Neurological Disorders Associated with Adult Cœliac Disease. By W. T. COOKE and W. THOMAS SMITH.

Vibratory Sense: A Critical Review. By DONALD B. CALNE and CHRISTOPHER A. PALLIS.

Polymyositis: A Survey of 89 Cases with Particular Reference to Treatment and Prognosis. By ARTHUR L. ROSE and JOHN N. WALTON.

Mental Deficiency Associated with Muscular Dystrophy—A Neuropathological Study. By N. PAUL ROSMAN and BYRON A. KAKULAS.

The Comparative Pathology of Cerebral Aneurysms. By M. R. CROMPTON.

The Pathogenesis of Cerebral Aneurysms. By M. R. CROMPTON.

Latent Sensory Aphasia in Hemisphere-Damaged Patients: An Experimental Study with the Token Test. By F. BOLLER and L. A. VIGNOLO.

***In Vitro* Examination of Phosphate Exchange between Plasma and Erythrocytes in the Blood of Multiple Sclerotic Patients.** By L. LATZKOVITS, I. HUSZÁK and F. SZÉCHENYI.

Notices of Recent Publications.

Index.

LONDON: MACMILLAN & CO. LTD

The British Journal of Educational Psychology

Editor L. B. BIRCH

Vol. 37, Part 1 — February 1967

H. LYTTON. Follow-up of an Experiment in Selection for Remedial Education.

D. F. SWIFT. Family Environment and 11+ Success: Some Basic Predictors.

D. S. FINLAYSON and L. COHEN. The Teacher's Role: A Comparative Study of the Conceptions of College of Education Students and Head Teachers.

S. MITCHELL and M. SHEPHERD. The Child Who Dislikes Going to School.

E. D. CAVE. The Effects of the 'Eleven-Plus' Result on the Subsequent Careers of Pupils.

M. L. PETERS. The Influence of Reading Methods on Spelling.

M. D. SHIPMAN. Environmental Influences on Response to Questionnaires.

E. G. RUMP and V. SOUTHGATE. Variables Affecting Aesthetic Appreciation in Relation to Age.

M. P. M. RICHARDS and H. E. ROSS. Developmental Changes in Children's Drawings.

H. J. EYSENCK. Intelligence Assessment: A Theoretical and Experimental Approach.

K. W. ROGERS. Concepts of Time in Secondary School Children of Above Average I.Q.

RESEARCH NOTES:

W. YULE, L. LOCKYER and A. NOONE. The Reliability and Validity of the Goodenough–Harris Drawing Test.

P. WILLIAMS. Regression Analysis Programmes for the Identification of Important Variables in Educational and Psychological Experiments.

R. J. GOLDMAN and D. F. CLARKE. The Minnesota Tests of Creative Thinking—A Note on Scorer Reliability in Follow-up Studies with English Primary School Children.

A. ABI RAFI. The Progressive Matrices (1938) and the Dominoes (D48) Tests: A Cross-Cultural Study.

R. R. DALE. Co-education as a Possible Influence on the Desire for Marriage.

J. K. BACKHOUSE. The Use of Valentine's Reasoning Tests in an Investigation into Transfer from Mathematics to Reasoning.

COMMENT:

H. LYTTON. Note on the Article 'Differences in Test Scores and in the Gaining of Selective Places for Scottish Children and those in England and Wales', by Douglas, Ross, Maxwell and Walker.

Etc., etc.

The Journal is published three times yearly. The subscription is £2 a year post free payable in advance, or £1 each number. Subscriptions and Orders should be sent to:

THE MANAGER, B.J.E.P. DEPARTMENT, METHUEN & CO. LTD
11 NEW FETTER LANE, LONDON, E.C.4

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 1–12
Printed in Great Britain

AN INFORMATION-PROCESSING EXPLANATION OF SOME PERCEPTUAL PHENOMENA*

By HERBERT A. SIMON
Carnegie Institute of Technology, Pittsburgh, Pennsylvania U.S.A.

An information-processing system that scans stimuli serially, part-by-part, and attempts 'simple' interpretations of the parts would experience a number of the well-known perceptual illusions that human subjects report. The hypothesized system has the same basic characteristics as systems which have been used to explain a wide range of cognitive phenomena. The description of the system is proposed as an explanation of some of the mechanisms for human perceptual processing.

I. Introduction

Contemporary schemes for using computers to process 'visual' two-dimensional and three-dimensional representations, along with the recent information-processing theories of human cognition provide significant new clues to a number of perceptual phenomena hitherto regarded as somewhat mysterious and inexplicable. This paper undertakes to show how some of the common visual 'illusions', and one auditory illusion, would arise in an information-processing system similar to those proposed recently to explain other aspects of cognition.

Newell, Shaw & Simon (1960) have advanced a theory of human problem-solving expressed in information-processing terms, Feigenbaum & Simon (1964) have proposed a theory of rote memory, and a number of investigators have proposed theories of concept attainment that are closely similar to one another (Hovland & Hunt, 1960; Laughery & Gregg, 1962; Feldman, 1963; Simon & Kotovsky, 1963, Johnson, 1964). While these theories have not been completely unified, they all share certain common assumptions. First, they assume that representations of stimuli and problem situations can be stored in memory in structures isomorphic with lists and branching list structures. These structures have the properties of associations, more precisely, of the 'directed associations' of the Würzburg School and Otto Selz. Secondly, the information-processing theories assume that active cognitive processes store and modify these structures, admitting communication between the central nervous system and the environment, and between short-term and long-term memory; and that the active processes are fundamentally *serial* processes which can handle only a small amount of material at a time (e.g. the 'seven chunks' of George Miller, 1956).

Information-processing theories have not progressed very far in accounting for sensation, but have been concerned largely with the activity of the central nervous system. Hence, using these theories to explain perceptual phenomena will be compatible with some recent approaches that emphasize the common parts of perception and cognition. Clearly, this is not the whole story of perception, but it may well be an important segment of it.

The reversibility of the Necker cube provides an introductory example for an information-processing explanation of a visual phenomenon. The 'impossible'

* This work was supported in part by grant MH-07722-01 from the National Institutes of Mental Health to the Carnegie Institute of Technology.

drawings for which the artist M. C. Escher is known make a second type of example, the phenomena of depth perception make a third example and apparent motion, visual and auditory, is discussed as a final example.

The basic approach is to show that almost any serial information-processing system that is capable of storing complex symbolic structures will 'perceive' stimuli, i.e. encode and store them, in such a way as to produce some of the phenomena that have been observed in human beings. In several cases, actual, running computer programmes serve as examples, even though these programmes were not written with simulation of human perception as a specific objective.

The explanations proposed appear to be consistent with a number that have been put forth in the recent literature on visual form recognition: for example, those of Gibson (1950) and Hochberg (1962). What distinguishes the present account is explicitness with respect to the character of the information processes needed by the viewer, and neutrality on the question of whether these processes are innate or acquired.

The Necker cube

The Necker cube provides a familiar demonstration of central control over visual perception. The subject is shown the two-dimensional diagram of Fig. 1 and asked to 'interpret' it as a three-dimensional cube. By asking him such questions as whether the square face on the lower left (*ABCD*), is in front of or behind the square face on the upper right (*EFGH*), it can be shown that two alternative interpretations are possible. Most subjects can learn to alternate 'voluntarily' (and will alternate involuntarily) between these two interpretations at a maximum rate of perhaps one alternation per second.

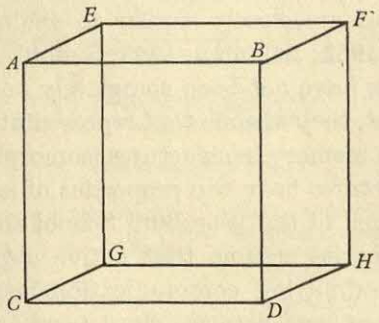


Fig. 1. The Necker cube.

From a subjective point of view, there is seldom any doubt in a subject's mind as to what is meant by *ABCD* being 'in front of', or 'behind' *EFGH*. When the subject is asked such a question, he can generally answer without hesitation. His report, however, is subject to all of the objections that can be raised against introspection in general. Is it reliable? How can its reliability be tested? And, operationally, what does it mean to say that the one square is 'in front of' or 'behind' the other? (As any fool can see, they both actually lie in the same plane!)

No answer is proposed to the question of reliability. That question has not often been raised about the Necker cube, possibly because the subjective reports are so readily produced, and because the experimenter himself, looking at the cube, is hard-

pressed to call them illusory. The operational meaning of the interpretation is another matter. To deny that the three-dimensional interpretation of the two-dimensional stimulus can be given operational meaning, would be to deny the whole possibility of an organism with two-dimensional sensory organs behaving 'as if' it lived in a three-dimensional world.

II. A Mechanism for interpreting visual stimuli

One way of providing operational meaning for the interpretation is to specify a mechanism for carrying it out. The mechanism might or might not resemble closely the actual human mechanism. It would show decisively how such a mechanism is possible, and could provide explanations for some of the phenomena in which this mechanism is implicated. Such a mechanism is to be specified.

A memory is postulated that is organized in terms of names of 'objects', and is capable of storing symbolized information about properties of objects and of relations between and among objects. In the case of the Necker cube, names of objects would include 'the point *C*', 'the angle *AEF*', 'the plane *AEBF*', 'the cube *ABCDEFGH*', and so on. Associated with the name '*C*', for example, might be information like: '*C* is a point', 'line *AC* joins points *A* and *C*', '*C* lies on the lower left-hand corner of square *ABCD*', '*GCD* is a right angle'. How information of this kind can be stored in a memory has just been illustrated: as English sentences on a piece of paper. It is easy to store the same information, in a wide variety of ways, in a digital computer.

One way in which the information could be stored in a computer is an analogical *image* of the diagram. In this phrase, the noun 'image' denotes that information inside the memory will not necessarily *state* (explicitly) the information about the stimulus, as it was stated in the English sentences above. Instead, the information in the memory will *represent* the stimulus, providing some of the information explicitly, some implicitly. The meaning of this distinction will become clear as we proceed. The image is described as 'analogical', to emphasize that the memory does not have to *be* three-dimensional (it does not contain a model of the cube), but has only to be capable of *representing* three dimensions symbolically.

A simple kind of analogical image for visual stimuli is provided by the coordinates of analytic geometry. A set of coordinates (x, y, z) can be associated with the name of each point in Fig. 1:

A	0, 1, 0	E	0, 1, 1
B	1, 1, 0	F	1, 1, 1
C	0, 0, 0	G	0, 0, 1
D	1, 0, 0	H	1, 0, 1

This representation of the cube contains, in fact, all of the information needed for a particular three-dimensional interpretation of the stimulus—the interpretation which places *EFGH*, as a vertical plane, behind the vertical plane *ABCD*. A little bit of the information is explicit: the locations of the points. Most of the information is implicit. For example, the fact that points *A*, *B*, *C* and *D* lie in the same vertical plane is deducible from the fact that they all have the same z-coordinate, and similarly for points *E*, *F*, *G* and *H*. The fact that the former plane lies in front of the latter plane is deducible from the fact that the z-coordinate of the former, 0, is smaller than

the z-coordinate, 1, of the latter. To deduce other facts—for example, that EGC is a right angle—requires slightly more complicated procedures.

A different set of coordinates could, of course, equally well have been assigned to the points, for example:

A	0, 1, 1	E	0, 1, 0
B	1, 1, 1	F	1, 1, 0
C	0, 0, 1	G	0, 0, 0
D	1, 0, 1	H	1, 0, 0

Applying the same deductive processes to the new co-ordinates, $ABCD$ and $EFGH$ are still vertical planes, but $ABCD$ now lies *behind* $EFGH$. In fact, the Necker cube has been reversed.

It is not suggested that the representations of visual stimuli in human memory are numerical or that information is extracted from them by doing analytic geometry. It is suggested that a representation analogical to the one just described, and with processes analogical to those of analytic geometry for interpreting the implicit information contained in the information, can account in general for human three-dimensional interpretation of two-dimensional stimuli, and in particular for the reversal of the Necker cube.

The explanation is not yet complete, howevever. A representation for three-dimensional objects has been provided that encodes such concepts as 'in front of', 'behind', 'right angle', and so on. But how the perceptual processes might transform a stimulus into this encoding has not been explained, nor how the two alternative encodings corresponding to the two ways of viewing the Necker cube might arise. Consider a process of the following kind:—

A. A scanner moves over the stimulus and detects simple configurations. By 'simple configurations' are meant figures like those the Gestalt psychologists call 'good': straight lines, especially if horizontal or vertical, right angles, squares, circles, closed symmetrical forms.

B. When a simple configuration has been detected, the scanner proceeds from it to the other elements of the stimulus, providing them with simple interpretations in relation to the initially discovered simple configuration. This process continues until an internal representation has been constructed for the entire stimulus, or a contradiction has been encountered.

C. If a contradiction is encountered, the interpretation is rejected and step *A* is repeated—usually with a different initial position of the scanner.

The example of the Necker cube enables the various elements of this scheme to be made more specific and more nearly operational. It is first assumed that a subject's physical orientation and environmental cues define for him the horizontal and vertical directions on the plane of the stimilus (so that he can notice that CD is horizontal and EG is vertical for instance). Suppose the scanner first encountered point F. It could then recognize the horizontal (EF) and vertical (FH) lines from F, and hence the right angle, $E\hat{F}H$. EF and FH, further, define a plane, which, since FH is vertical, and EF horizontal, is interpreted as a vertical plane, parallel to the surface of the paper.

Following the closed line, $FEGHF$, in this plane, the square $EFGH$ is detected. Returning to point F, the line FB must be accounted for. The angles it makes with EF and FH are not right angles. Line FB can be given a 'simple' interpretation, however, by interpreting B as lying behind the plane of square $EFGH$. The 'simple' interpretation of $BFDH$, then, is that it is a vertical square behind the plane of the

paper and perpendicular to it. (*B* could equally well have been interpreted as *lying in front of* the plane of *EFGH*. Whichever interpretation is taken, provided that it is taken consistently, the Necker cube phenomena will appear. Empirically, it seems that subjects generally interpret such points as lying behind the previously defined plane.) A continuation of the same processes finally leads to the cube with *EFGH* as its front face and *ABCD* as its rear face.

If exactly the same processes had been applied, starting with point *C* instead of point *F*, the stimulus would again have been interpreted as a cube, but with *ABCD* as its front face and *EFGH* as its rear face. Thus the alternation of cubes can be produced by starting the scanner alternately at points *C* and *F* (or points *B* and *G*). It is well known that this is actually one procedure that subjects consciously use to produce rapid alternation of the two interpretations.

The scheme described, while adequate to account for the reversibility of the Necker cube would undoubtedly run into difficulty in simulating the way human beings interpret complex visual displays. It has been developed in detail to illustrate the kinds of information processes that are involved. A scheme capable of far more complex and realistic performances, but based on the same general principles, has been developed and implemented on an electronic computer by Roberts (1963) and by Kirsch (1964), to the last of which the reader is referred for a detailed description.

Even Roberts' (1963) scheme, as he was careful to point out, fell short of using all the perceptual cues that humans have been shown, by Gibson (1950) and others, to employ. For example, it does not use texture to estimate distance. Addition of these cues to the scheme is entirely feasible, and would represent no change in the point of view. The question of specifically what cues the human viewer uses, and the extent to which these are innate or learned, can be separated from the question of whether the human viewer is fundamentally the kind of serial information processor that has been described.

'Impossible' figures

The proposed explanation can be tested further by applying it to other phenomena of the same general kind as alternation of the Necker cube. As a preliminary, it is important to explain what is meant by 'contradictory' interpretations (see step *C* of the interpretive scheme). Suppose that, for some reason, scanning of the Necker cube had begun independently at *both* points *C* and *F*. Then an interpretation might have been arrived at (see Fig. 2), in which both *ABCD* and *BDFH* might be interpreted as the front and right side of one cube. A subject can, in fact, rather easily train himself to 'see' the figure in precisely this way. (This possibility is predicted by the theory proposed here and, as far as the writer is aware, has not previously been noticed.) As long as the figure is interpreted as a set of four planes, perpendicular in pairs, the interpretation causes no difficulty. The moment, however, that the scanner seeks to interpret, say, the edge *AC* in relation to these planes, a contradiction arises. For *AC*, as an edge of square *ACEG*, lies *behind* the surface plane, while *AC*, as an edge of *ABCD*, lies *in* the surface plane. If the scanner focuses on *AC*, it cannot maintain the interpretation of the two intersecting cubes. The interpretation is readily maintained by alternating attention between, say, *EG* and *BD*, each of which has a single, unambiguous location in the interpretation.

The phenomena of consistency and contradiction may be described thus. The scanning process operates locally, gradually building up an interpretation of the stimulus, and extending this interpretation to adjoining parts. If the same part is reached by two different routes, the two interpretations it thus acquires may be identical or different. In the latter case, a perceptual contradiction has occurred, and the original interpretation that led to it will be maintained with difficulty; more precisely, it will only be maintained by restricting the scanning to the non-contradictory portions of the stimulus.

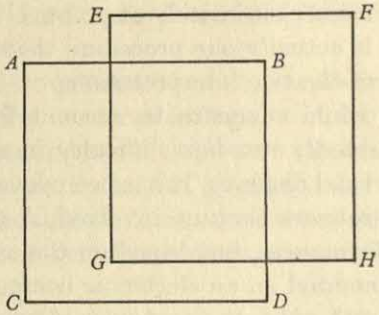


Fig. 2. Partial Necker cube.

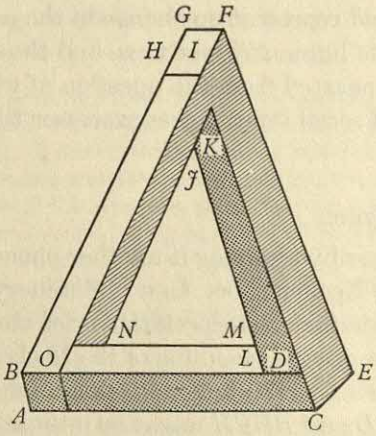


Fig. 3. 'Impossible' figure.

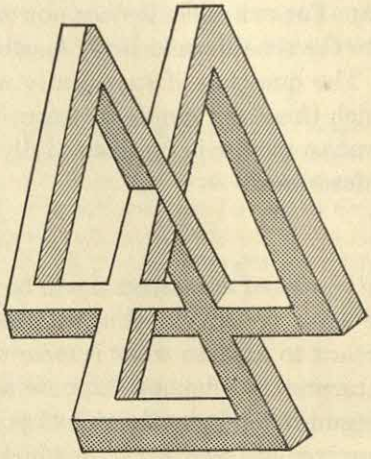


Fig. 4. 'Impossible' figure.

This explanation of the perception of 'contradiction' applies also to the 'impossible' figures that have been constructed by the painter M. C. Escher, and by Penrose & Penrose (1958) and others. (See Figs. 3 and 4.) If the scanner starts at A in Fig. 3, for example, it might arrive at an interpretation of ABC and DKL as lying in one plane (say, a vertical plane), BGF and BML as lying in a second (horizontal) plane perpendicular to the first, and CEF and IKN as lying in a third plane, perpendicular to the other two. But an attempt to assign coordinates to the points on this basis would lead immediately to contradictory assignments, for the points O and I would both have to lie at the intersection of all three planes, which is patently impossible.

The 'plausibility' of these figures, in the sense of the immediate reaction that they can be interpreted as three-dimensional, arises from the fact that they are locally interpretable. The contradiction arises from the fact that no local interpretation can be extended to the figure as a whole. The scanner finds itself in the position of a surveyor who, after going around the three sides of a triangle discovers that the sum of the angles does not equal 180°. He is advised, thereby, that the figure he has been measuring cannot be interpreted as a triangle in the Euclidean plane.

The explanation for the interpretation of two-dimensional stimuli as representing three-dimensional objects, and for the concept of contradiction can readily be applied to familiar 'illusions', like those produced by the Ames rooms. The main property

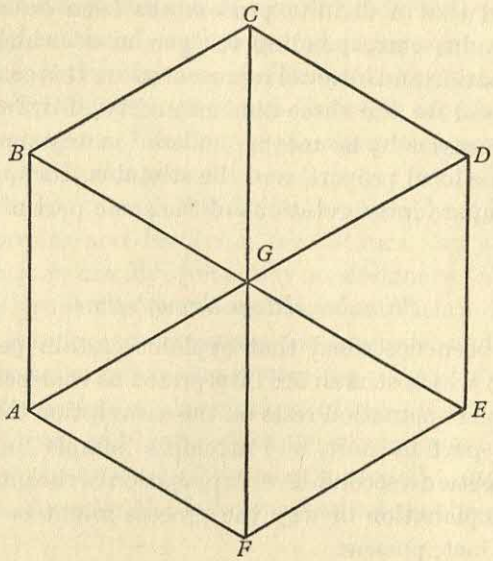


Fig. 5. Special case of Necker cube.

required of the scanner to produce these illusions is a strong tendency to interpret angles as right angles, wherever possible. In a photograph of an Ames room showing two persons of apparently quite different sizes, the illusions can, with effort, be 'reversed' by focusing attention on the pair of human figures, and ignoring the walls of the room until the figures appear to be of the same actual size, but at different distances. If this is done, the angles in the room will appear to lose some of their rectilinearity. In fact, the mechanisms that have been described seem generally adequate, with little amplification, to handle all the perceptual phenomena discussed by Ittelson & Kilpatrick (1952).

Two-dimensional and three-dimensional interpretations

The stimulus of Fig. 5 is more often than not interpreted by subjects as a two-dimensional regular hexagon. It can also, of course, be interpreted as a three-dimensional cube with front face *ABGF*, or a cube with front face *CDEG*. The theory should explain why most subjects arrive first at the former, two-dimensional interpretation rather than one of the three-dimensional interpretations.

The scanner, starting at A would discover the vertical line AB, but no horizontal line. Further, it would discover the line AGD extending across the figure. It would discover that each of the three lines, CF, AD, and BE are axes of bilateral symmetry, and that the six triangles about G are equi-lateral and equal. Hence, the figure, interpreted as two-dimensional, would be composed of symmetrically disposed simple sub-figures. Any attempt to reinterpret the stimulus in three dimensions would destroy a number of these properties. The six triangles no longer exist in the three-dimensional figure; the halves of the six diameters become distinct line segments, parallel but not colinear.

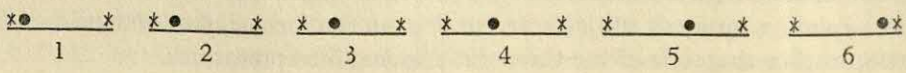
The foregoing explanation is not unlike those commonly given by Gestalt theorists. What distinguishes it from them is that operational meaning has been given to notions like 'good figure', and that a definite process has been described for discovering properties in the stimulus corresponding to 'goodness' which have been used to construct the interpretation and internal representation. It is consistent with measures that have been proposed for the three-dimensionality of drawings (cf. Hochberg & Brooks, 1960). The process is by no means 'holistic' in any simple sense. It proceeds sequentially to examine local properties of the stimulus. Its apparent 'holism' arises from its ability to compare interpretations of the same part of a stimulus arrived at by alternative paths.

Stimulus change through time

A mechanism has been described that explains certain perceptual phenomena, specifically the way in which stimuli are interpreted as representing objects in three-dimensional space. The explanation rests on the assumption that a scanner examines the stimulus serially, part by part, and attempts 'simple' interpretations of these parts. In the present section a second-level explanation of the interpretive process itself is proposed, i.e. an explanation of why the process might be expected to have the properties it does, in fact, possess.

The organism's perceptual processes may be conceived as one part of an information storage and retrieval system. Suppose that a subject is given the laboratory task of observing a stimulus, reporting what he has observed; observing a second, reporting about it; and so on. Or, to change the situation slightly, but significantly, suppose he is given the task of observing a series of stimuli and, after observing all of them, reporting what he has seen. With both tasks, the subject must have some means of storing what he has seen until he reports. In the second task, since the reporting is delayed, efficiency in the storage and retrieval processes becomes critically important. How can efficiency be increased?

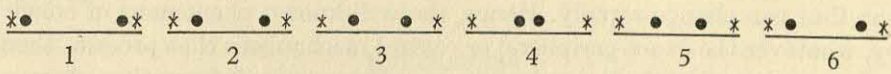
The means available for increasing efficiency in information processing depend on the nature of the successive stimuli. Redundancy from one stimulus to the next is the principal means of reducing storage requirements. One form of redundancy occurs when parts of successive stimuli can be set in correspondence with each other that is, when they can be interpreted as belonging to the 'same object'. Consider the following series of stimuli:



A relatively succinct interpretation is that a dot moves from the left-hand to the right-hand side of an interval. An information-processing system could accomplish this interpretation as follows: In the first stimulus, names are created for 'the dot' and 'the interval'. The stimulus might be described by: 'the dot is at the left end of the interval'. Alternatively, the interval might be divided into locations, numbered from 1 to 6, from left to right, and the stimulus could be described by: 'the dot is at location 1'. The entire sequence of stimuli is then simply: 'the dot moves from location 1 to location 6 at a constant rate'. A slightly more formal translation would be '$\mathrm{Loc}(d, t) = t$'. As before, it is not being proposed that the description that is stored is actually verbalized in any of these ways, simply that what is stored might be roughly isomorphic with one of these verbalizations.

The central device used here to capitalize on the redundancy is that of treating the stimuli as though they consisted of objects and changes in, or actions upon or by objects, or of relations among objects (cf. Michotte, 1946). Linguists are almost in agreement that virtually all languages have noun-like words and verb-like words-names and processes or actions. The noun–verb distinction has essentially been used in describing the stimulus in the present case.

One can imagine a world in which this distinction would not be very useful, a world of clouds constantly forming and dissolving, for instance. Our world is quite different. It is a world in which it is usually quite easy to designate 'objects' that, properly described, retain their properties over considerable stretches of time, and change in simple, orderly ways, moving continuously, for example, and in smooth paths. In such a world, an information processing system can exploit the redundancy that is present by storing information as descriptions of the objects it has identified and named. This storage scheme will both reduce the storage required and, what is often more important, reduce the amount of processing needed to update the information and to keep it current. Consider the following variant on the previous stimulus:



In this case the stimulus can be interpreted as two dots, one moving from right to left, the other from left to right; alternatively, it can be interpreted as two dots moving toward the middle of the interval, colliding, and rebounding. With either interpretation, the identities of the two dots are preserved throughout.

The usefulness of this scheme, in our world, requires effective procedures for identifying objects, and effective procedures for describing them 'properly', i.e. in ways that will leave their properties most nearly invariant over time. The problem of proper description is considered first.

Suppose that the stimulus is a cube of wire. If it is held in a fixed position, it is equivalent, in stationary monocular vision, to a two-dimensional drawing of a Necker cube. Let it be rotated, however, and, as is well known, it becomes vividly three-dimensional. It is obvious that, under rotation, three-dimensional representation of the cube is far more parsimonious than two-dimensional, for the three-dimensional figure can be described as an invariant cube, subject to a simple rotation.

In kinematics it is shown that a rigid body has only six degrees of freedom. No

matter how irregular the shape of the body, nor how complex its motion, that motion, and the totality of information about the body, can be described by giving the time paths of six coordinates. Hence, in a world like ours, description of a stimulus as a collection of three-dimensional rigid bodies will generally be much more parsimonious than description of the same objects in two dimensions.

With the three-dimensional representation, objects in visual space will generally have a well-defined shape, which changes slowly if at all, and a well-defined size, which also tends to be invariant through time. Of course, illusions and difficulties in interpretation can be created by introducing contradictory clues, as has been noted previously; but such clues will seldom be present in nature.

The scheme proposed requires that the information processing system should have means for identifying objects. Motion is again exploited, by peripheral as well as by central processes, for motion generally provides more reliable cues than are provided by colour tone or hue contrasts. In modern painting the Impressionists, in particular, exploited the unreliability of the cues of tone and hue to 'dissolve' the world of real objects into a television screen that was supposed to represent the 'real' impact on the retina. (The Impressionist approach, and pointillism in particular, nevertheless, incorporated some erroneous views about perception. In particular, the Impressionists were unaware, apparently, of the peripheral encoding of cues derived from movement to determine object boundaries; cf. Gombrich, 1960.)

In a world of rigid, moving objects, where cues from movement are more important than cues of tone and hue in fixing object boundaries, how would an information-processing system be expected to record colour information? One important generalization is that 'colour', would be expected to be associated with each object, to remain relatively constant over the object and to remain relatively invariant under movement and changing conditions of light and shade. This colour might, of course, contain cues as to the position of the object in space, the 'blueing' of distant objects for instance, but one would expect it to be rather independent of those aspects of lighting that can change rapidly. Hence, the well-known phenomena of colour constancy, whatever the exact peripheral or central mechanisms that produce them, are generally consistent with the principle of efficiency in information storage and retrieval (cf. Helson, 1943).

Shepard (1964) has recently demonstrated an amusing and striking auditory illusion that can be interpreted in a parallel manner to the foregoing interpretation of apparent movement. His demonstration adds further credence to the hypothesis that such illusions arise centrally rather than peripherally, and that they are strongly moulded by requirements of parsimony. The familiar 'barber's pole' illusion fits the earlier discussion of the apparent motion of dots along a line. Shepard (1964) has constructed an auditory 'barber's pole' by presenting a mixture of tones and by gradually increasing the pitch of each component of the mixture while fading out the higher components and introducing new components of lower pitch. The hearer perceives the sound as continually rising in pitch although, after it has 'risen' a full octave, the mixture of tones is again identical with the initial mixture. The perceptive reader will note that the phenomenon has been described as it might be by a listener who was taken in by the illusion, for to speak of 'each component' of the mixture as gradually rising in pitch, is to give particular components at one moment

the same name as the corresponding components at the next moment. But this naming procedure is not intrinsic to the physical sounds; it is imposed by the listener —it is precisely what creates the illusion.

III. Conclusion

This paper has shown how a wide range of the known phenomena of visual perception can be accounted for by a mechanism that encodes information from the stimulus into an internal representation having certain efficient characteristics from an information processing and retrieval standpoint. A detailed account has been given of the Necker cube and related phenomena, and the theory has been extended in more general terms to other perceptual phenomena, including constancy of size, shape, and colour.

In many respects the proposed theory is far from novel. It is anticipated by many discussions of perceptual phenomena by Gestalt theorists and others. What is perhaps new is the demonstration that a small number of particular, operationally definable mechanisms can be specified that will produce many of the known effects. Realization of these mechanisms on a computer and simulation of the perceptual behaviour, as has been begun by Kirsch (1964) and others, would provide evidence as to whether these mechanisms perform, qualitatively and quantitatively, in a way that is consistent with the known phenomena.

I am indebted for many of the ideas, incorporated in these schemes to notions derived from B. F. Green's Baseball program; from Robert K. Lindsay's 'Inferential memory as the basis of machines which understand natural language' (1963); from my own work on the Heuristic Compiler; and from work of Thomas G. Evans of Massachusetts Institute of Technology, e.g. 'A heuristic program to solve geometry–analogy problems' (1964). My interest in applying these ideas to perception was stimulated by conversations with Jacob Beck, and I have benefited from comments from him and from Profs. E. G. Boring, E. J. Gibson and Julian Hochberg. Needless to say, my gratitude for their help does not imply their agreement with the theory presented here.

References

Evans, T. G. (1964). A heuristic program to solve geometry–analogy problems. *Proc.* 1964 *Spring Joint Computer Conf., Am. Fed. Information Processing Societies Conf. Proc.* **25**, 327–38.

Feigenbaum, E. A. & Simon, H. A. (1964). An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *J. verb. Learn. verb. Behav.* **3**, 385–96.

Feldman, J. (1963). Simulation of behaviour in the binary choice experiment. In *Computers and Thought*, ed. E. A. Feigenbaum and J. Feldman. New York: McGraw-Hill.

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton-Mifflin.

Gombrich, E. H. (1960). *Art and Illusion.* New York: Pantheon Books.

Helson, H. (1943). Some factors and implications of color constancy. *J. opt. Soc. Am.* **33**, 555–67.

Hochberg, J. (1962). The psychophysics of pictorial perception. *Audio-vis. Commun. Rev.* **10**, 22–52.

Hochberg, J. & Brooks, V. (1960). The psychophysics of form: reversible-perspective drawings of spatial objects. *Am. J. Psychol.* **73**, 337–54.

Hovland, C. I. & Hunt, E. B. (1960). The computer simulation of concept attainment. *Behavl. Sci.* **5**, 265–7.

Ittelson, W. H. & Kilpatrick, F. P. (1952). Experiments in perception. *Scient. Am.* **185**, 50–5.

JOHNSON, E. S. (1964). An information-processing model of one kind of problem solving. *Psychol. Monogr.* no. 581.

KIRSCH, R. A. (1964). Computer interpretation of English language text and picture patterns. *IEEE Trans. on Electronic Computers*, EC-13, 363–76.

LAUGHERY, K. W. & GREGG, L. W. (1962). Simulation of human problem solving behavior. *Psychometrika* **27**, 265.

LINDSAY, R. K. (1963). Inferential memory as the basis of machines which understand natural language. In *Computers and Thought*, pp. 217–36, ed. E. A. Feigenbaum and J. Feldman. New York: McGraw-Hill.

MICHOTTE, A. (1946). *La Perception de la Causalité*. Paris. [English translation by T. R. Miles & E. Miles (1963). *Perception of Causality*. New York: Basic Books.]

MILLER, G. A. (1956). The magical number seven. *Psychol. Rev.* **63**, 81.

NEWELL, A., SHAW, J. C. & SIMON, H. A. (1960). Report on a general problem-solving programme. *Proc. Int. Conf. Information Processing, Paris, June* 1959. London: Butterworths.

PENROSE, L. S. & PENROSE, R. (1958). Impossible objects: a special type of visual illusion. *Br. J. Psychol.* **49**, 31–3.

ROBERTS, L. G. (1963). *Machine Perception of Three-Dimensional Solids*. Technical Report no. 315, Lincoln Laboratory. Cambridge: Massachusetts Institute of Technology.

SHEPARD, R. N. (1964). Circularity on judgments of relative pitch. *J. acoust. Soc. Am.* **36**, 2346–53.

SIMON, H. A. & KOTOVSKY, K. (1963). Human acquisition of concepts for sequential patterns. *Psychol. Rev.* **70**, 534–46.

(*Manuscript received* 29 *June* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 13–27
Printed in Great Britain

SENSORY SPATIAL AFTER-EFFECTS IN RELATION TO CHRONOLOGICAL AGE, MENTAL RETARDATION AND SCHIZOPHRENIA

By R. H. DAY, R. BURNS,
Monash University
G. SINGER, V. HOLMES and D. LETCHER
University of Sydney, Australia

Claims that the visual and kinaesthetic spatial after-effect vary as a function of chronological age, mental retardation, and schizophrenia have been tested in four experiments. The principal after-effect studied was that involving judgements of the horizontal following stimulation by a slanted pattern or object. In Exps. III and IV more traditional visual (alignment) and kinaesthetic (width estimation) tasks were also investigated. Data from Exp. I show that no variation in the kinaesthetic after-effect occurs with age but suggest that the visual after-effect may decline with age. The latter may be because adjustment times increase with age. The data from Exps. II and III indicate that there is no difference between normal and mentally retarded subjects in two visual and two kinaesthetic after-effects. Experiment IV showed that schizophrenic subjects do not differ from normal and psychopathological controls in two visual and two kinaesthetic after-effects.

After stimulation of a sensory system involved in spatial discrimination (vision, hearing, and the tactile, kinaesthetic, and vestibular senses), judgements of the spatial and spatio-temporal properties of the stimulus undergo change compared with pre-stimulation judgements. These modifications of spatial judgement have been referred to as negative (Gibson, 1933), figural (Köhler, 1940) and spatial (Day & Singer, 1964) after-effects, and where movement is involved, as the after-effect of movement.

During the last two decades considerable effort has been devoted to establishing the main determinants of these effects, to their relation to other sensory phenomena, and to their explanation (see reviews by Day, Pollack & Seagrim, 1959; McEwen 1958; Spitz, 1958). Determinants of the effect can conveniently be divided into two classes: those deriving from properties of the stimulus (e.g. intensity, spatial pattern, temporal features) and those associated with characteristics of the individual. Since Freeman & Josey (1949) claimed a relation between the after-effect of visual motion and impairment of memory, and Klein & Krech (1952) claimed an association between figural after-effect and cortical damage, there has been increasing interest in the second class. The purpose of the experiments reported here was to investigate systematically under standard conditions of equipment and procedure the effects of chronological age, intellectual level and schizophrenia on the magnitude of the sensory spatial after-effect in vision and kinaesthesis.

In a series of recent experiments (Day & Singer, 1964), large, easily measurable and consistent after-effects were obtained in both vision and kinaesthetis. The task was the relatively simple one of judging the apparent horizontal before and after stimulation (for 90 sec) by a slanted pattern. Preliminary work showed that this task could be carried out adequately by both children and the mentally retarded; hence it was possible to use this task and procedure for all groups.

Apparatus and general procedure

Since the main apparatus and general procedure were essentially similar for all groups they are described first. Apparatus used in subsidiary experiments with mentally retarded and schizophrenic subjects is described in the relevant sections.

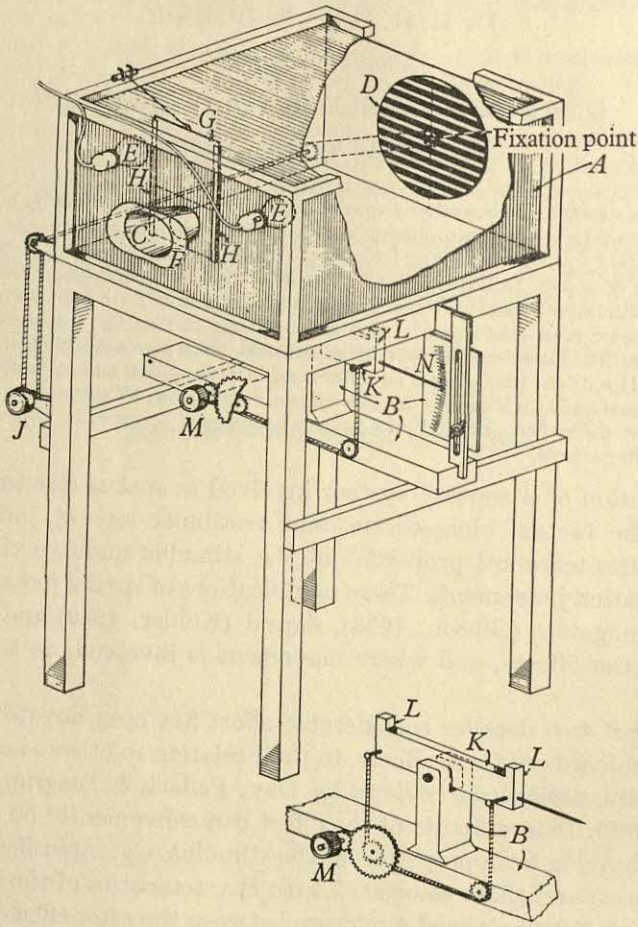


Fig. 1. Apparatus used in Expts. I, II and IV for the measurement of a visual and kinaesthetic after-effect. A, light-tight box; B, kinaesthetic component; C, viewing aperture and facepiece; D, aluminium disk; E, lamps; F, sponge rubber facepiece; G, metal slide; H, guides for metal slide; J, visual control, K, kinaesthetic bar; L, stops 12 in. apart on kinaesthetic bar; M, kinaesthetic control; N, scale indicating angular displacement of kinaesthetic bar.

Apparatus

Visual component. The apparatus for the measurement of the visual spatial after-effect shown in Fig. 1 consisted of a light-tight box 38 cm high, 61 cm wide and 76 cm deep, mounted so that a viewing aperture and facepiece at one end were at approximate eye-height for a seated subject. At the opposite end of the box, and inside it, was a 28 cm flat black aluminium disk with 13 mm bars painted with 13 mm interspaces in luminescent paint and a black 7 mm fixation point at its centre. This disk was mounted so that it rotated about a central axis, and so that its slant from the true horizontal could be read from a scale on the outside of the box at the far end from the subject. Angular readings could be made to the nearest 0·25°. Inside the box and above

the viewing aperture were two 100 W clear filament lamps which were used at regular intervals to activate the luminescent bars on the disk so that they glowed green in the dark. The viewing aperture was fitted with a moulded sponge rubber facepiece which served to exclude all exterior light from the box when the subjects face was pressed into it. A metal slide under spring tension could be raised or lowered quickly from the top of the box to expose or to occlude the bars. A control knob permitted the subject to orientate the bars to any angular position.

Kinaesthetic component. The kinaesthetic spatial after-effect was measured by means of an apparatus (Fig. 1) consisting of a 30 cm × 7 cm bar pivoted about is centre and so placed that it could be comfortably reached by the subject's outstretched hand. Two stops at either end defined a 30 cm excursion of the subject's hand across its length. A control on the subject's left was used to adjust the slant of the bar into any angular position about its centre and a protractor scale permitted angular readings to the nearest 0·25°. To control the frequency of hand motion across the bar a metronome set throughout at 84 beats/min was placed out of the subject's view. In all kinaesthetic trials a pair of ski-goggles in which the lenses were covered with opaque paper were used to blindfold the subject.

General procedure

Throughout the experiments in which these two pieces of equipment were used a trial consisted of four pretests, stimulation with the stimulus object slanted 15° from the horizontal for 90 sec, and a single post-test. In the pretests and post-tests the subject was always required to adjust the stimulus object so that it seemed horizontal. The after-effect was the difference between the mean of the pretests and the post-test in degrees. The starting position of the stimulus objects for pre- and post-tests varied randomly between 0° (horizontal) and 2·5° and 5°, up and down on the subjects right. Each subject underwent two trials under any one condition; one with the bar slanted 15° up on the subject's right and the other down. The score for any condition was the subject's average after-effect for the two trials and all statistical treatments are based on this score.

An after-effect in an expected direction was scored as positive and that in an unexpected direction as negative. Thus, if the bar was slanted 15° down on the subject's right during kinaesthetic stimulation, it would be expected that the physically horizontal bar would be judged as down on the left in the post-test; to compensate for this the subject would be expected to slant the bar down on the right for it to appear horizontal. This adjustment was scored as positive if it was further down on the right than in the pretest.

Subjects were encouraged to make adjustments as quickly but as accurately as possible and, for reasons which are made clear below, all post-test adjustments were timed with a stop-watch. In the visual after-effect trials subjects were instructed to fixate the point at the centre of the disk. Since, however, the disk subtended a visual angle of approximately 17°, losses of fixations would not have been as critical as for patterns of smaller subtense. When the eye was fixated on points removed from the fixation point substantial areas of the retina would nevertheless have been stimulated by the pattern of slanted bars.

After practice in making visual and kinaesthetic adjustments to the horizontal from the various starting positions, each subject underwent four trials, two for the kinaesthetic (K) after-effect and two for the visual (V). In one trial of each pair the stimulus object was slanted 15° up (U) on the subject's right, and in the other 15° down (D) during the stimulation period. Hence, there were four conditions: KU, KD, VU, VD. In order to counterbalance order effects, these conditions were arranged in a 4 × 4 Latin square to which subjects were assigned randomly.

Effects of chronological age: Experiment I

Köhler & Wallach (1944) invoked the notion of 'permanent satiation' and used it to account for generally small displacements of contours which they observed following stimulation. Permanent satiation, they argued, developed progressively and cumulatively over extended time-spans. It could be expected, therefore, that young children in whom the permanently satiated state had less time in which to develop would be more sensitive to stimulation and thus exhibit greater figural after-effects.

Köhler & Dinnerstein (1947) suggested, however, that there would be a low positive correlation between the size of the effect and age, a prediction that seems contrary to expectations from the 'permanent satiation' hypothesis. Pollack (1960) investigated the frequency of occurrence of two visual after-effects using subjects whose ages ranged from 5 years to adult. He used a relatively brief stimulation period (10 sec) and expected that the after-effects would occur only in the case of subjects particularly prone to them. The data demonstrated a decline in frequency of occurrence between 4 and about 11 years, followed thereafter by an increase to adult age. A second investigation was made by Thurner & Seyfried (1962) of the kinaesthetic after-effect in relation to age using a width estimation task similar to that used originally by Köhler & Dinnerstein (1947). They found no difference between 12-year-old children and adults in the size of the effect, its diminution over time, the frequency of the opposite effect, or the goodness of fit of the data to a logarithmic function. The discrepancy between the data reported by Pollack (1960) and by Thurner & Seyfried (1962) is not inconsistent with the observation that visual and kinaesthetic after-effects are not significantly correlated (Spitz & Lipman, 1960).

Singer & Day (1965) in an investigation of the dissipation function for the kinaesthetic after-effect found that it was exponential with the most rapid decline occurring in the first few seconds after cessation of stimulation. It was found further that younger subjects took longer to make an adjustment than adults. This result suggests that differences in the measured magnitude (as opposed to the true magnitude) of the after-effect may in some instances be attributable to its greater dissipation during the post-test adjustment period. For this reason post-test adjustment times were recorded throughout.

Method

Subjects

There were 175 subjects from primary and high schools in seven age-groups of twenty-five subjects each. The age-groups were 7–8, 9–10, 11–12, 13–14, 15–16, 17–18 and 19–20 years. All subjects of school age were in the grade normal for their age-group.

Apparatus and procedure

The apparatus and procedures for the measurement of both the visual and kinaesthetic after-effect were as described above.

Results

The means and S.D. for the visual and kinaesthetic after-effects, and for the adjustment times under the two conditions, are shown in Table 1. Each mean is based on twenty-five subjects and on two measurements per subject.

A separate analysis of variance for the visual and the kinaesthetic after-effect data has shown that whereas the visual after-effect varies significantly with age ($F = 14{\cdot}04$; D.F. 6, 168; $P < 0{\cdot}01$), the kinaesthetic after-effect does not ($F = 1{\cdot}55$; D.F. 6, 168; $P < 0{\cdot}1$). The correlation between the magnitude of the after-effect and the post-test adjustment time over all subjects is significantly negative for the visual effect ($r = -0{\cdot}26$; $P < 0{\cdot}01$) and for the kinaesthetic effect ($r = -0{\cdot}17$; $P < 0{\cdot}05$). These significantly negative correlation coefficients show that with decrease in the magnitude of the after-effect there was an increase in the time taken for subjects to make the post-test adjustment. Analyses of variance of the adjustment time data in relation

to age has shown that there are significant differences in the cases of both the visual task ($F = 2{\cdot}70$; D.F. 6, 168; $P < 0{\cdot}05$) and the kinaesthetic task ($F = 2{\cdot}42$; D.F. 6, 168; $P < 0{\cdot}05$).

The product-moment correlation coefficient between the visual and kinaesthetic after-effect for all 175 subjects just failed to achieve significance ($r = +0{\cdot}142$; $0{\cdot}1 > P > 0{\cdot}05$).

Table 1. *Mean visual and kinaesthetic spatial after-effects and post-test adjustment times and* S.D. *for seven age-groups in Expt. I* ($n = 25/group$)

		Age-groups (yr)						
		7–8	9–10	11–12	13–14	15–16	17–18	19–20
Visual after-effect (degrees)	Mean	+ 3·20	+ 2·35	+ 2·40	+ 2·27	+ 1·46	+ 1·57	+ 1·62
	S.D.	1·82	1·16	1·59	1·13	0·31	0·23	1·00
Adjustment time (sec.)	Mean	7·50	7·98	8·38	10·44	10·44	10·50	9·18
	S.D.	2·47	4·19	3·48	4·69	4·28	5·86	3·34
Kinaesthetic after-effect (degrees)	Mean	+ 4·12	+ 4·98	+ 5·09	+ 4·25	+ 4·13	+ 4·10	+ 3·62
	S.D.	2·60	3·10	2·02	1·56	1·93	1·27	1·66
Adjustment time (sec.)	Mean	15·66	16·42	14·88	18·18	18·42	21·96	19·86
	S.D.	6·73	6·35	6·67	6·80	9·39	11·15	8·36

Effects of mental retardation: Experiments II and III

The cortical field (satiation) theory of the sensory spatial after-effect has led to investigations of the effect in relation to states often attributed to malfunctioning of the C.N.S. The rationale of this view is that, if the cortical mechanisms associated with these states are different from those of normal individuals modifiability of the brain which is said to underlie spatial after-effects may also vary. Accordingly the spatial after-effects of mentally retarded and brain-injured individuals have been compared with those of normal controls. It has been predicted that both states involve poorer capacity to 'satiate' with after-effects of smaller magnitude in consequence (Prysiazniuk & Kelm, 1963; Spitz & Blackman, 1959; Spitz & Lipman, 1961).

The empirical data reported so far are by no means unequivocal. Klein & Krech (1952) found that for brain-injured subjects the kinaesthetic spatial after-effect occurred more frequently, was of greater intensity, achieved a maximum more rapidly, and persisted for a longer period. Jaffe (1954) found that there was no difference between brain-injured and normal subjects for the same after-effect. Spitz & Lipman (1961) reported that mentally retarded subjects exhibited significantly smaller kinaesthetic after-effects and concluded that retarded subjects have 'weaker' satiation processes than normal subjects of the same chronological age. Prysiazniuk & Kelm (1963) and Spitz & Blackman (1959), using different procedures, reported visual after-effects of greater magnitude for matched normal controls than for mentally retarded children. In neither study, however, is there any indication that the after-effects were significantly different from zero. In point of fact Prysiazniuk and Kelm reported after-effects in the non-expected direction for two of the test periods with the mentally retarded subjects. There was also a significant difference

between the two groups in adjustment times which, as pointed out above, could be a determinant of differences in the measured after-effects.

Experiment II

In view of the discrepancies between predictions of the magnitude of spatial after-effects for normal and mentally retarded subjects and differences between independent experimental data, the aim of this experiment was to compare the magnitude of the visual and kinaesthetic spatial after-effects of mentally retarded and normal children of the same chronological and of the same mental ages. Because of possible variations in post-test adjustment times between the three groups, these times were also measured for all groups.

Method

Subjects

There were three groups each of thirty-six subjects, details of whose chronological and mental ages and I.Q.s (Terman–Merrill, 1937) are shown in Table 2. The retarded group were pupils in a special school for 'slow learners' and the normal controls were drawn from primary and high schools in the same area.

Apparatus and procedure

The apparatus and procedures were again those that have already been described above.

Table 2. *Mean chronological ages,* I.Q. *and mental ages, and age and* I.Q. *ranges of the three groups for Expt. II* ($n = 36/group$)

	Groups		
	Mentally retarded subjects	Normal subjects of same chronological age	Normal subjects of same mental age
Chronological age			
Mean	14 years 6 months	14 years 5 months	9 years 8 months
Range	13 years 1 month to 15 years 3 months	14 years 1 month to 15 years	9 years to 10 years 3 months
I.Q.			
Mean	69·2	119·0	116·3
Range	59–87	104–135	102–132
Mental age			
Mean	9 years 8 months	17 years 3 months	11 years 9 months

Results

Means and S.D. for visual and kinaesthetic after-effects in degrees of arc and for the corresponding adjustment times are shown for the three groups in Table 3. Single classification analyses of variance (Winer, 1962) were calculated separately for the visual and kinaesthetic after-effect data. The differences between the visual after-effects for the three groups are not significant ($F = 1{\cdot}03$; D.F. 2, 105; $P > 0{\cdot}05$). For the kinaesthetic data the mean after-effect for the retarded groups was greater than for the normal groups of the same chronological age and only slightly less than for normal subjects of the same mental age. An analysis of these data showed that the differences between the groups are significant ($F = 4{\cdot}56$; D.F. 2, 105; $P < 0{\cdot}05$). 'Protected' t tests (Winer, 1962) between the three groups, however, have shown that no pair of differences achieves significance ($P > 0{\cdot}05$), but that both the mean visual and kinaesthetic after-effects differ significantly from zero ($P < 0{\cdot}001$). Of the 216

visual after-effects, 11 % were in a non-expected direction and so were 6 % of the kinaesthetic after-effects. Analyses of variance have shown that there was no significant difference in adjustment times between the two groups for either the visual or the kinaesthetic tasks. The correlation between the visual and the kinaesthetic after-effects based on all subjects was +0·005 (N.S.).

Table 3. *Mean visual and kinaesthetic spatial after-effects and post-test adjustment times and* S.D. *for three groups in Expt. II* (n = 36/*group*)

		Groups		
		Mentally retarded subjects	Normal subjects of same chronological age	Normal subjects of same mental age
Visual after-effect (degrees)	Mean	+ 2·16	+ 1·28	+ 2·70
	S.D.	2·10	1·05	1·13
Adjustment time (sec)	Mean	9·95	9·92	7·34
	S.D.	6·29	4·68	4·00
Kinaesthetic after-effect (degrees)	Mean	+ 5·12	+ 3·87	+ 5·65
	S.D.	3·32	1·94	2·34
Adjustment time (sec)	Mean	15·08	16·76	15·87
	S.D.	6·35	7·39	7·66

Experiment III

It is conceivable that the discrepancy between the results from Expt. II and those found in previous investigations may derive from the different methods used to generate and to measure the visual and kinaesthetic after-effects. In the experiments of Prysiazniuk & Kelm (1963) and Spitz & Blackman (1959) the visual after-effect involved the discrimination of alignment and of size respectively. The experiment by Spitz & Lipman (1961) on the kinaesthetic after-effect involved the widely used width estimation task originally described by Köhler & Dinnerstein (1947). A possible source of difference between the results from this and earlier studies is the stimulus conditions used to generate and measure the visual after-effect. The stimulus figures used by Prysiazniuk & Kelm (1963) and by Spitz & Blackman (1959) were such that slight changes in visual fixation could have markedly affected the magnitude of the after-effect. In Expt. II above, however, the effects of loss or change of fixation would have been less profound, since the inducing and test patterns were spatially coincident in terms of total area (the test pattern fell within the area of the inducing pattern). For these reasons it seemed worth while to compare the visual and kinaesthetic after-effects of mentally retarded and normal control subjects using similar tasks to those used earlier by these other workers. It was considered desirable to do so in the interests of completeness and in order to establish whether the discrimination tasks used to study spatial after-effects constitute a critical variable in investigations of this type.

Method

Apparatus

The apparatus for the induction and measurement of a visual after-effect resembled that used by Prysiazniuk & Kelm (1963) and is shown in Fig. 2. It consisted of two pieces of 76 × 76 cm Plexiglass mounted on a wooden frame and arranged so that the nearer piece, which was hinged

below, could be quickly folded down to reveal the piece beneath. A red fixation point, 1 mm in diameter, was located at the centre of the first Plexiglass square (Fig. 2A). The two elements of the inducing figure were located 5 mm above and below the fixation point and 7 mm to its left and right. Each rectangular element was 19 mm × 28 mm and was cut from matt black tape. The black test figure on the farther Plexiglass square (Fig. 2) consisted of two 3 × 18 mm rectangles, one above the other vertically and 14 mm apart. Midway between them was a red fixation point similar to and coincident with that on the inspection figure. The lower half of this

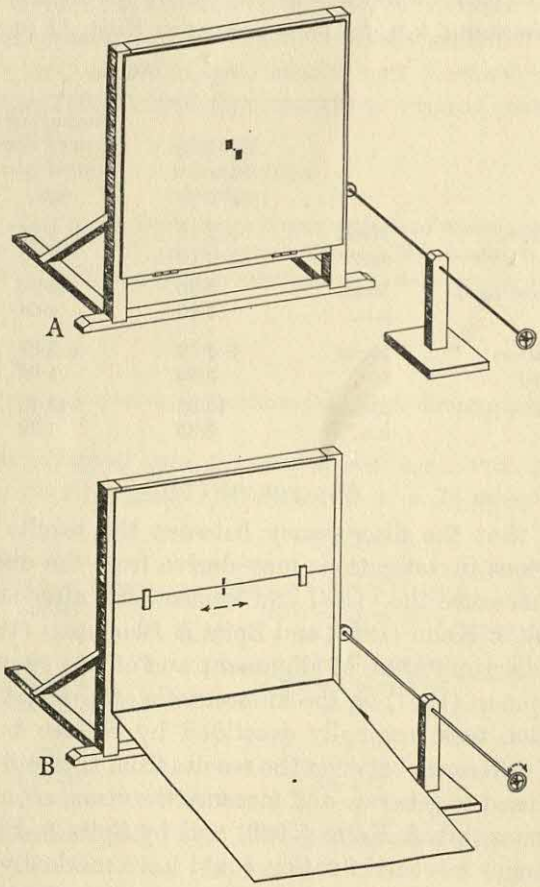


Fig. 2. Apparatus used in Expts. III and IV for measurement of a visual after-effect. A, Inducing pattern; B, test pattern with arrows indicating the direction of movement of the lower component.

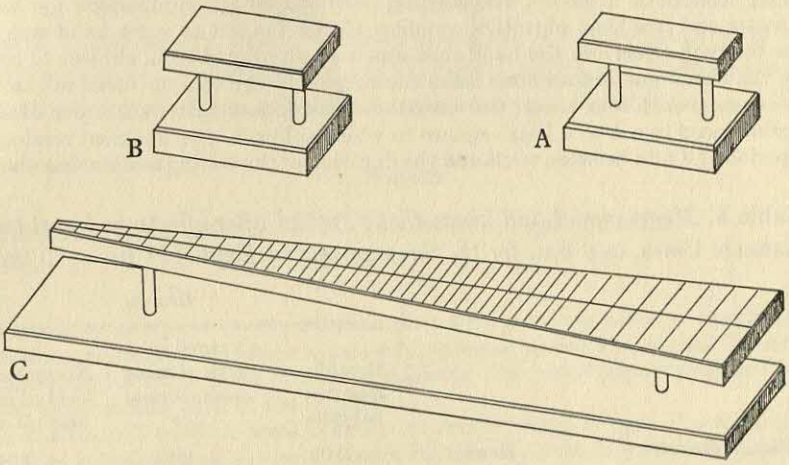
Plexiglass square could be moved from left to right by means of a control on the subject's right. During the pretest and post-test the subject was required to adjust the position of the lower test element so that it appeared to be aligned vertically with the upper element. This adjustment was achieved by moving the lower half of the Plexiglass by means of the rotary control. A scale to the supporting frame permitted readings of misalignment to the nearest 0·25 mm.

The apparatus for the induction and measurement of the kinaesthetic after-effect shown in Fig. 3 was the same as that used by Spitz & Lipman (1961) which in turn was similar to that used by Klein & Krech (1952). The inducing object (Fig. 3A) was a wooden block 15 × 6·3 × 1·9 cm mounted on a supporting base. The test object (Fig. 3B) was 3·8 cm wide and otherwise similar to the inspection object. The tapered comparison object (Fig. 3C) was 76 cm long, and was 10 cm wide at one end and 1·3 cm at the other. In its other dimensions it was the same as the inducing and test objects and was similarly mounted. A centimetre scale attached along the

length of the comparison object permitted measurements to the nearest 0·5 mm when a ruler was placed across the width of the scale at the position of the tips of thumb and forefinger. The test and inducing bars were placed on one table and the comparison bar on another. During a trial the subject stood between the two tables with the wider end of the comparison bar behind him.

Subjects

There were three groups of twenty subjects each from the same schools and grades as in Expt. II. Details of chronological and mental ages and I.Q.s (Terman–Merrill, 1937) are shown in Table 4.



BUREAU OF EDUCATIONAL & PSYCHOLOGICAL RESEARCH
LIBRARY
GOVT. OF WEST BENGAL
CALCUTTA

Fig. 3. Apparatus used in Expts. III and IV for measurement of a kinaesthetic after-effect. A, Test object; B, inducing object; C, comparison object.

Table 4. *Mean chronological ages,* I.Q. *and mental ages, and age and* I.Q. *ranges of the three groups for Expt. III* ($n = 20$/*group*)

	Groups		
	Mentally retarded subjects	Normal subjects of same chronological age	Normal subjects of same mental age
Chronological age			
Mean	14 years 0 months	14 years 11 months	9 years 10 months
Range	13 years 1 month to 14 years 10 months	14 years 5 months to 15 years 8 months	9 years 4 months to 11 years 6 months
I.Q.			
Mean	70·6	112·3	105
Range	57–79	100–122	95–122
Mental age			
Mean	9 years 11 month	16 years 10 months	10 years 9 months

Procedure

A trial consisted of four pretests, stimulation for 90 sec, and a single post-test. For the visual task the subject was required to align vertically the two elements of the test figure during pretests and post-test using the rotary control. During the visual stimulation period the subject was instructed to fixate the red point in the inspection figure; at the end of the period the Plexiglass was lowered quickly and the adjustment of the test figure elements to apparent alignment was made. The difference between the mean pretest adjustments and the post-test adjustment in millimetres was the measure of the after-effect. This difference was again designated as positive or negative according to whether it was in an expected or non-expected direction for a visual after-

Bureau of Ednl. & Psyl. Research
(S. C. E. R. T.)
Date .....
Acc. 5.897 .....

effect. The starting position of the lower element of the test figure varied randomly between 0·5 mm to left and right of the upper element. Each subject underwent two trials, one with the starting position on the right and the other with it on the left.

During the pre-tests and post-tests of the kinaesthetic task, the subject was required to grasp the test bar with the thumb and forefinger of one hand and to match its width with the same fingers of the other hand on the tapered comparison bar. During stimulation the thumb and forefinger of the hand used on the test bar were moved back and forth along the edges of the 6·3 cm wide inducing bar and, on completion of this period, the test bar was quickly placed in position for the post-test. The difference between the mean position of the hand on the comparison bar for the pretests, and its position during the post-test served as a measure of the after-effect and was scored as positive or negative. The starting positions for the comparison bar were varied between greater and less than objective equality. Either the left or right hand was placed on the test bar for both trials and the hand that was used alternated from subject to subject.

For both the visual and kinaesthetic tasks the subject's score was the mean of the two trials. Each subject underwent four trials, two visual and two kinaesthetic; the order of these trials was counterbalanced in a 4 × 4 Latin square to which subjects were assigned randomly. There was a rest period of 3 min between trials and the duration of the testing session was about 30 min.

Table 5. *Mean visual and kinaesthetic spatial after-effects and post-test adjustment times, and* S.D. *for the three groups in Expt. III* ($n = 20/group$)

		Groups		
		Mentally retarded subjects	Normal subjects of same chronological age	Normal subjects of same mental age
Visual after-effect (degrees)	Mean	+ 0·03	+ 0·02	+ 0·04
	S.D.	0·04	0·02	0·03
Adjustment time (sec.)	Mean	15·30	15·80	15·25
	S.D.	6·87	4·51	4·97
Kinaesthetic after-effect (degrees)	Mean	+ 1·16	+ 2·65	+ 1·01
	S.D.	5·72	2·84	4·66
Adjustment time (sec.)	Mean	12·27	13·25	11·70
	S.D.	5·56	4·04	3·20

Results

The means and S.D. for the visual and kinaesthetic after-effects, and for the corresponding adjustment times are shown in Table 5. Separate analyses of variance for the visual and kinaesthetic scores have shown that there is no significant variation in the magnitude of either effect between the three groups $F < 1$ for both visual and kinaesthetic after-effects; but in each case the mean for all three groups combined is significantly different from zero ($P < 0·01$). Unlike the data from Expt. II however, 38·3 % of the 120 visual after-effects and 37·6 % of the 120 kinaesthetic after-effects were in a non-expected direction. There were no significant differences in adjustment times between the three groups for either task ($F < 1$). The correlation between the visual and kinaesthetic after-effects based on all subjects was $r = +0·13$ (N.S., $P > 0·05$).

EFFECTS OF SCHIZOPHRENIA: EXPERIMENT IV

The satiation theory of the spatial after-effect (Köhler & Wallach, 1944) was invoked by Wertheimer (1954) to predict that schizophrenic subjects would show smaller after-effects than normal individuals. It was argued that because of inefficiency

in metabolism the schizophrenic would show less satiation after stimulation and therefore smaller after-effects. Experimental results reported by Wertheimer (1954), Wertheimer & Jackson (1957) and Kelm (1962) have confirmed this prediction with the additional but unexpected finding that schizophrenics exhibit significant effects in a non-expected direction. This latter result was noted by Kelm (1962) and is implicit in the data of Wertheimer (1954). Two of these studies can be questioned on methodological grounds. Wertheimer (1954) failed to use pretests so that it is difficult to establish whether in fact stimulation resulted in a significant after-effect. Kelm (1962) did not take into account the direction of change between pre-tests and post-tests in comparing groups and used instead the absolute magnitude of change. The fourth experiment in this series was therefore carried out to compare the visual and kinaesthetic after-effects of schizophrenic and normal subjects.

Method

Apparatus

The apparatus was that described above and shown in Figs. 1, 2 and 3.

Subjects

There were three groups of twenty subjects each. One group consisted of hospitalized schizophrenic patients, another of hospitalized non-schizophrenic patients to control for factors due to hospitalization, and a third group of normal subjects. The latter group consisted of members of the psychiatric nursing staff of the hospital.

Selection criteria. All subjects were between 17 and 37 years of age. The schizophrenic and non-schizophrenic hospitalized subjects did not include any with known brain-damage or with evidence of brain-damage in their case-histories. Also excluded were patients admitted less than 5 days before with a diagnosis of acute 'brain syndrome', patients who had undergone electro-convulsive therapy in the previous 6 months, and alcoholic and drug-addicted patients. The schizophrenic group included only patients diagnosed as having schizophrenia and with no other type of diagnosis. None was in remission. The non-schizophrenic hospitalized group included diagnoses of any psychotic disorder other than schizophrenia, but excluded patients with hallucinations, delusions or gross thought disorders. Most patients were under medication but those who appeared grossly affected by it were not tested at that time. Several testing sessions were deferred for this reason. All subjects in the two pathological groups agreed to take part in the experiments.

Procedure

Each subject underwent two trials on each of four tasks. These were the visual and kinaesthetic tasks of Expt. I and of Expt. III. The procedures were exactly as they have been described above for Expts. I and III. Each subject underwent all four tasks within a single session, with a 10 min rest between two 20 min testing periods.

Results

The mean visual and kinaesthetic after-effects and S.D. are shown in Table 6 for the three groups. Separate analyses of variance have been computed for each of the four tasks since the scores for each are different in kind. The results of these analyses are summarized in Table 7. No F ratio achieved an acceptable level of significance ($P > 0{\cdot}1$). Thus for each of the four after-effect tasks there was no significant difference between schizophrenic, non-schizophrenic hospitalized and normal subjects.

The adjustment times for the four tasks are also shown in Table 6 and the analyses of variance of these data in Table 8. For all four types of after-effect there is a

Table 6. *Means of two visual and two kinaesthetic spatial after-effects and post-test adjustment times, and* S.D. *for the three groups in Expt. IV* ($n = 20/group$)

		Groups		
		Schizophrenic patients	Non-schizophrenic hospitalized patients	Normal subjects
Visual horizontal task after-effect (degrees)	Mean	+ 2·80	+ 2·16	+ 2·64
	S.D.	1·43	1·17	1·04
Adjustment time (sec.)	Mean	6·37	3·70	4·50
	S.D.	2·41	1·50	1·22
Visual alignment task after-effect (mm)	Mean	− 0·37	− 0·51	− 0·71
	S.D.	0·49	0·45	0·36
Adjustment time (sec.)	Mean	11·45	8·59	9·15
	S.D.	4·67	2·92	2·16
Kinaesthetic horizontal task after-effect (degrees)	Mean	+ 4·84	+ 3·80	+ 4·62
	S.D.	1·88	2·21	1·46
Adjustment time (sec.)	Mean	13·81	9·86	7·55
	S.D.	1·69	5·29	1·65
Kinaesthetic width estimation task after-effect	Mean	+ 2·50	+ 3·25	+ 3·30
	S.D.	12·80	15·80	13·25
Adjustment time (sec.)	Mean	10·77	6·20	6·35
	S.D.	3·29	2·28	2·19

Table 7. *Analyses of variance of the two visual and the two kinaesthetic after-effect scores for the three groups in Expt. IV*

Condition	Source	S.S.	D.F.	M.S.	F	
Visual horizontal task	Between	4·60	2	2·30	1·44	N.S.
	Within	90·89	57	1·59		
Kinaesthetic horizontal task	Between	11·90	2	5·95	1·56	N.S.
	Within	216·40	57	3·79		
Visual alignment task	Between	1·56	2	0·78	2·37	N.S.
	Within	18·83	57	0·33		
Kinaesthetic with estimation task	Between	0·32	2	0·16	0·01	N.S.
	Within	496·02	57	0·7		

Table 8. *Analyses of variance of post-test adjustment times for the three groups in Expt. IV*

Condition	Source	S.S.	D.F.	M.S.	F	P
Visual horizontal task	Between	77·08	2	38·54	11·47	< 0·001
	Within	191·76	57	3·36		
Kinaesthetic horizontal task	Between	481·18	2	240·59	8·19	< 0·001
	Within	1673·58	57	29·36		
Visual alignment task	Between	91·98	2	45·99	3·72	< 0·05
	Within	703·22	57	12·34		
Kinaesthetic width estimation task	Between	270·24	2	135·12	18·53	< 0·001
	Within	415·49	57	7·29		

significant difference between the three groups owing to the longer adjustment times for the schizophrenic groups (cf. Table 6).

The percentages of after-effects in a non-expected direction for the four tasks are in agreement with those found in Expts. II and III. For the visual and kinaesthetic tasks involving judgements of the horizontal, these percentages were 5·8 % and 5·0 % respectively; they were 44·1 % and 33·3 % for the visual alignment and kinaesthetic width estimation tasks. All the after-effects are significantly different from zero (cf. Table 6), but in the case of the visual alignment task the effect was in the opposite direction to that expected.

Discussion

Contrary to predictions derived from the notion of permanent satiation (Köhler & Wallach, 1944), the data from Expt. I indicate that there is no change in the magnitude of the kinaesthetic spatial after-effect as a function of chronological age. This finding is in agreement with that of Thurner & Seyfried (1962) who used a different inducing and testing procedure. Adjustment times for the kinaesthetic task increased with age and correlated negatively with the after-effect. On the other hand the visual after-effect decreased with age whereas adjustment times again increased. There was a significant negative correlation between the visual after-effect and visual adjustment times. Although at first sight these results appear inconsistent, this is not necessarily the case. First, in confirmation of Spitz & Lipman (1960), there was no significant correlation between kinaesthetic and visual spatial after-effects. Secondly, since adjustment times for the visual after-effect are in the range 7·5–10·9 sec (cf. Table 1), during which rapid dissipation of the effect occurs (Hammer, 1949; Oyama, 1953), even slight changes in time would affect the after-effect. By contrast, adjustment times for the kinaesthetic after-effect are in the range 14·8–21·9 sec (cf. Table 1), during which the after-effect dissipates relatively slowly (Singer & Day, 1965). It is not surprising then that although significant correlations occur between after-effect and adjustment times for the visual and kinaesthetic tasks, and significant increases in adjustment times with age for both, only the visual after-effect varies with age. Even so, the visual data do not permit a firm conclusion. Since both visual after-effect and visual adjustment times vary with age and are correlated, the former may be a function of the latter or the after-effect may in fact vary alone. Further enquiry will be necessary to establish this point.

Previous findings have indicated significantly smaller visual and kinaesthetic after-effects for mentally retarded subjects compared with those for normal controls. In Expt. II and III no such difference occurred between groups although a significant after-effect occurred for each group. Of interest is the observation that the visual and kinaesthetic after-effect in judgements of the horizontal yielded more consistent effects, with fewer non-expected directions of the after-effect than the more commonly used tasks. Whereas the percentages of non-expected directions of after-effect for the main task were 11 % (visual) and 6 % (kinaesthetic), those for the more traditional methods were 38 % (for visual alignment) and 37 % (for kinaesthetic width estimation). These findings suggest that the induction and measurement of visual and kinaesthetic spatial after-effects with mentally retarded individuals can be achieved more effectively with tasks involving judgements of the horizontal after stimulation by a slanted object.

No significant difference was found between schizophrenic, non-schizophrenic hospitalized and normal subjects for the two visual and the kinaesthetic after-effects in Expt. IV. These data are generally inconsistent with earlier results for schizophrenic subjects, but the frequency with which non-expected directions of after-effect occurred for the commonly used tasks may be the explanation for the inconsistency. For the visual alignment task 49 % of the 120 after-effects that were measured were in a non-expected direction compared with 5·8 % for the horizontal judgement task. In the kinaesthetic width-judgement task 33 % of responses were opposite to the expected direction compared with 5 % for the horizontal judgement task. The higher frequency of reversed effects for the alignment and width estimation tasks calls into question the reliability of earlier findings and the occurrence of 'opposite' effects for schizophrenic subjects (Kelm, 1962).

Two methodological issues have emerged, in addition to the generally negative findings from the experiments reported here so far as age, mental retardation and schizophrenia are concerned. First, there is evidence from the experiments concerned with the effects of chronological age and schizophrenia to suggest that post-test adjustment times should be controlled in experiments of this type. Since both the visual spatial after-effect and adjustment times varied with age it is not possible to establish whether the former varies with age or is a function of the latter. Secondly, there has been a tendency to induce and to measure visual and kinaesthetic after-effects using a restricted range of techniques involving size discrimination. This investigation indicates that these traditional methods result in a higher frequency of non-expected directions of the effects. The principal methods involving judgements of the horizontal used here exhibited significantly fewer such reversals and provide a simple and effective means of induction and measurement.

REFERENCES

DAY, R. H. & SINGER, G. (1964). Spatial after-effects within and between kinesthesis and vision. *J. exp. Psychol.* **68**, 337–43.

DAY, R. H., POLLACK, R. H. & SEAGRIM, G. N. (1959). Figural after-effects: a critical review. *Austral. J. Psychol.* **11**, 14–45.

FREEMAN, E. & JOSEY, W. E. (1949). Quantitative visual index to memory impairment. *Arch. neurol. Psychiat.* **62**, 794–7.

GIBSON, J. J. (1933). Adaptation, after-effect, and contrast in the perception of curved lines. *J. exp. Psychol.* **16**, 1–31.

HAMMER, E. R. (1949). Temporal factors in figural after-effects. *Am. J. Psychol.* **62**, 337–54.

JAFFE, R. (1954). Kinesthetic after-effects following cerebral lesions. *Am. J. Psychol.* **67**, 668–76.

KELM, H. (1962). The figural after-effect in schizophrenic patients. *J. nerv. ment. Dis.* **135**, 338–45

KLEIN, G. S. & KRECH, D. (1952). Cortical conductivity in the brain injured. *J. Pers.* **21**, 118–48.

KÖHLER, W. (1940). *Dynamics in Psychology*. New York: Liverwright.

KÖHLER, W. & DINNERSTEIN, D. (1947). Figural after-effects in kinesthetics. In *Albert Michotte: Miscellanea Psychologica*, pp. 196–220. Louvain: Catholic University of Louvain.

KÖHLER, W. & WALLACH, H. (1944). Figural after-effects: an investigation of visual processes. *Proc. Am. phil. Soc.* **88**, 269–357.

MCEWEN, P. (1958). Figural after-effects. *Br. J. Psychol. Monogr. Suppl.* no. 31.

OSGOOD, C. E. & HEYER, A. W. (1951). A new interpretation of figural after-effects. *Psychol. Rev.* **59**, 98–118.

OYAMA, T. (1953). Experimental studies of figural after-effects: (1) temporal factors. *Jap. J. Psychol.* **23**, 239–54.

POLLACK, R. H. (1960). Figural after-effects as a function of age. *Acta Psychol.* **17**, 417–23.

POLLACK, R. H. (1962). A note on figural after-effects and age. *Acta Psychol.* **20**, 252–3.

PRYSIANIUK, A. W. & KELM, H. (1963). Visual figural after-effects in retarded adults. *J. abn. soc. Psychol.* **67**, 505–9.

SINGER, G. & DAY, R. H. (1965). Temporal determinants of a kinaesthetic after-effect. *J. exp. Psychol.* **69**, 343–8.

SPITZ, H. H. (1958). The present status of the Köhler–Wallach theory of satiation. *Psychol. Bull.* **55**, 1–28.

SPITZ, H. H. & BLACKMAN, L. S. (1959). A comparison of mental retardates and normals on visual figural after-effects and reversible figures. *J. abn. soc. Psychol.* **58**, 105–10.

SPITZ, H. H. & LIPMAN, R. S. (1960). The reliability and intercorrelation of individual differences on visual and kinaesthetic figural after-effects. *Percept. mot. skills.* **10**, 159–66.

SPITZ, H. H. & LIPMAN, R. S. (1961). A comparison of mental retardates and normals on kinaesthetic figural after-effects. *J. abn. soc. Psychol.* **62**, 686–7.

THURNER, F. K. & SEYFRIED, H. (1962). Are figural after-effects dependent upon age? *Acta Psychol.* **20**, 58–68.

WERTHEIMER, M. (1954). The differential satiability of schizophrenic and normal subjects: a test of a deduction from the theory of figural after-effects. *J. gen. Psychol.* **51**, 291–9.

WERTHEIMER, M. (1955), Figural after-effects as a measure of metabolic efficiency. *J. Pers.* **24** 56–73.

WERTHEIMER, M. & JACKSON, C. W. (1957). Figural after-effects, 'brain modifiability', and schizophrenia: a further study. *J. gen. Psychol.* **57**, 45–54.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

(*Manuscript received* 18 *April* 1966)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 29–38
Printed in Great Britain

ADAPTING WITKIN'S FIELD INDEPENDENCE THEORY TO ACCOMMODATE NEW INFORMATION FROM AFRICA

By MALLORY WOBER

*Department of Social Anthropology, University of Edinburgh**

Theories developed in America hold that the same cognitive style informs both the pattern of an individual's social relations and his performance in certain tests. The tests used in America have largely provided problems in a visual idiom. It is hypothesized that cultures found in Africa may attune skills relatively more highly in an auditory or proprioceptive, than in the visual sphere dominant in Euro-American cultures. Results using a visual, and a mixed visual and proprioceptive test suggest a different pattern of response from African than from American subjects. This is taken to show that the sameness of cognitive style through all fields of an individual's behaviour may not occur for Africans as it may for Americans. Differences of cognitive style for proprioceptive problems may have their own correlates of social behaviour patterns, and these remain to be studied.

A modern rapprochement between theories of personality, and of intellectual development, has been made by Witkin and his collaborators who introduced the concept of field-independence (Witkin, Lewis, Hertzman, Machover, Bretnall-Meissner & Wapner, 1954). Receiving support from the work of Bennett (1956), Epstein (1957), Young (1959) and others, though with some equivocal findings as for example from Gruen (1955), the position was thoroughly reviewed and re-stated in terms of the concept of psychological differentiation by Witkin, Dyk, Faterson, Goodenough & Karp (1962).

Recent work in Africa is relevant to refining Witkin's theories, and it is suggested that the skills involved in the performance of tests used by Witkin need to be more carefully understood than hitherto. Particular skills might be associated with particular cultural backgrounds, and the finding of a given degree of psychological differentiation might not be so generalized throughout all aspects of an individual's functioning as Witkin's theories suggest. Use is made of the idea of 'sensotypes' put forward by Wober (1966) to suggest that in the field of visual perception and the exercise of allied powers of psychological differentiation, Witkin's schema associating social and psychological test data may hold good; but that it remains to be shown if it also holds good for transactions in the field of proprioception, which may be culturally of salient importance in parts of Africa.

Theory of field-independence and psychological differentiation

Field-independence was defined by Witkin *et al.* (1962, p. 47) as the ability to separate an item perceived from its context. Most experiments used visually presented material which was described in the context of the 'visual field'. An extension of the theory linked visual with social and maturational phenomena. The ability to separate visual items from their context was shown to be aligned with the development of a sense of distinct personal identity. The person is considered to be an item, set in the context or 'social field' of his family and the society around him.

* Now at King's College, Cambridge.

Witkin *et al.* (1962) describe two extremes of practices in the socializing of children. The first tends to produce a conformist child through strict punishment and by preventing the assumption of individual responsibility. Such children are said to 'perceive globally'; they do not structure the perceived world actively and hence they tend not to deal skilfully with problems of visual analysis. The other extreme produces the individualist, who is considered to have to assess the world for himself; because he develops the habit of analysing structure, he is able to do well at certain types of problem. In this way, the theory becomes a theory of psychological differentiation: tests of field independence, social conformity and general ability all measure activities which can be described in similar terms.

Dawson (1963) was the first to apply field independence theory to findings in Africa. He noted that socialization practices in certain African tribal societies resembled the situation producing the classical extreme of Witkin's field-dependent type. People were taught to respond to social standards, to be 'other-directed' in Riesman, Glazer & Denney's (1953) nomenclature rather than individualistic. Further, tests of field-independence of the kind used by Witkin evoked very poor performance from Dawson's subjects (Africans from the Temne tribal group), suggesting a lack of habits of analysis for some types of visual material.

Two kinds of reason were sugggested for these poor performances of Temne subjects. One was Witkin's orthodox explanation, that a high degree of socialization as distinct from individualization underlay a low level of psychological differentiation which was responsible both for poor test performance and for the patterns of social behaviour observed. The second explanation concentrated on the lack of familiarity which Temne subjects had with Western modes of communication, especially print, diagrams and transactions in visual codes; this lack of familiarity would explain poor test scores, and socialization practices and social behaviour would remain to be accounted for separately. This second explanation, however, received no support from the observations of Preston (1964) and the experiment of Berry (1965). They showed that Eskimos, as unlettered as Temne and with equally little contact with Western cultures, did much better on tests such as Block Designs than did Temne. The findings left the orthodox Witkin–Dawson explanation of Temne socialization practices and resultant psychological differentiation in a strengthened position.

This was the stage at which the present refinement to the above theory was introduced. It was noticed that the tests used by Witkin of field-independence and of psychological differentiation involved transactions and skills in the visual field. Certain tests also involved the subjects in integrating data from visual and proprioceptive fields, though the term 'field' appears to have been used by Witkin *et al.* (1962) ambiguously to refer to sources of both modalities of perception. It appeared possible that American subjects, brought up in a culture emphasizing visual values and codes for communicating information, might be a valid population in which the Witkin theory could stand. However, African (Temne) subjects were brought up in a culture emphasizing proprioceptive values and skills, and thus their relevant field of psychological differentiation might be that of proprioception or audition and, at any rate, not so much the field of visual transactions. Evidence on the nature of Temne and other West African cultures includes the absence of writing previous to Western or Muslim contact, the shortage of colour terms, and even of analysis of stars into

constellations by pattern. On the other hand, Temne and other West African cultures include an elaboration of the proprioceptively and aurally perceived world. Thus, music is an extension of speech, rhythm an extension of movement, beauty a function of grace of movement as much as of configuration of visage, and dance is a regular and favoured form of elaboration of activity, started at an early age.

The idea of sensotypes has therefore been put forward by Wober (1966), that is, that the prevailing patterns of childhood intake and proliferation of information from the various sense modalities differ according to culture. Thus, while Witkin's whole work was done among Americans largely of one sensotype, the application of his theories and tests might not be valid for Africans of another sensotype. The use of one of Witkin's tests involving proprioception might therefore be a more appropriate index of psychological differentiation among Africans of the type tested by Dawson (1963) and Berry (1965) than the use of the entirely visual tests.

The following hypotheses need to be tested as a preliminary to future work on the question of the appropriate sensory modality for tests of differentiation:

(1) That in the sensotype common in certain West African cultures, proprioceptivity is (still) relatively more elaborated with respect to visuality than would be the case in Western cultures;

(2) That, therefore, West African subjects would have better scores on tests where proprioceptivity was important: (*a*) relative to their scores on tests dependent on visuality; and (*b*) relative to scores of Westerners on similar tests;

(3) That results of West African subjects would approximate more towards those of any Western groups in whom proprioceptivity was highly trained than to results of normal Western groups.

Method

Two tests of the group described by Witkin *et al.* (1962) as being indices of field-independence were included in a battery given to subjects in Southern Nigeria. The mean scores of the Nigerian subjects have been compared with mean scores found by American workers, and correlations between test scores and an index of educational level have been examined for the Nigerian subjects.

The tests

The Rod and Frame Test (*RFT*). The subject is seated in a totally dark room and views a display consisting of a luminous square frame and a luminous rod mounted centrally within it. The frame has its bottom edge parallel to the ground, but it can be tilted about a horizontal axis so that either of the bottom corners can be pointing almost towards the ground. The rod is mounted on the same axis and can be pointed towards the ground ('vertical') or tilted away from the vertical with or against the tilt of the square. In the apparatus used, the dark room was a 6 ft. 6 in cube internally. The luminous square had sides 13 in. long, and the width of the slit actually visible was ¼ in. The rod was 11 in. long and also bore a slit ¼ in. wide. An arrangement was made with bulbs, a rheostat and a battery charger, whereby just enough current was supplied to make the slits comfortably visible.

The subject deals not only with a visual display, but in a sense also with a proprioceptive display, as he sits in a high-backed armchair which is tiltable to right or left by wedges placed under the legs. A footboard is attached to the chair, which keeps the subject's feet off the ground. Both the subject in his chair, and the square frame and rod may therefore be tilted (see Fig. 1).

The task for the subject is to tell the experimenter to move the rod which has been tilted away from the vertical before the illumination is made, back to a vertical position. When the subject believes the rod is resting in a vertical position, its actual displacement from the vertical is read

off by an assistant outside the dark room, who views a suitably arranged pointer and scale. This error from the true vertical constitutes the score. A high score thus indicates a poor performance.

The Embedded Figures Test (*EFT*). The subject is shown a simple geometrical figure on a card. He then sees a complex figure which contains the original simple figure as a part of it (the first card having been removed). The task is for the subject to find where the simple figure is 'hidden' and then to point it out. The present version of the test used eight items from those of Witkin (1950). These items were A-3, A-4, C-1, C-4, D-1, D-2, F-1 and G-1 (see Fig. 2). The time taken

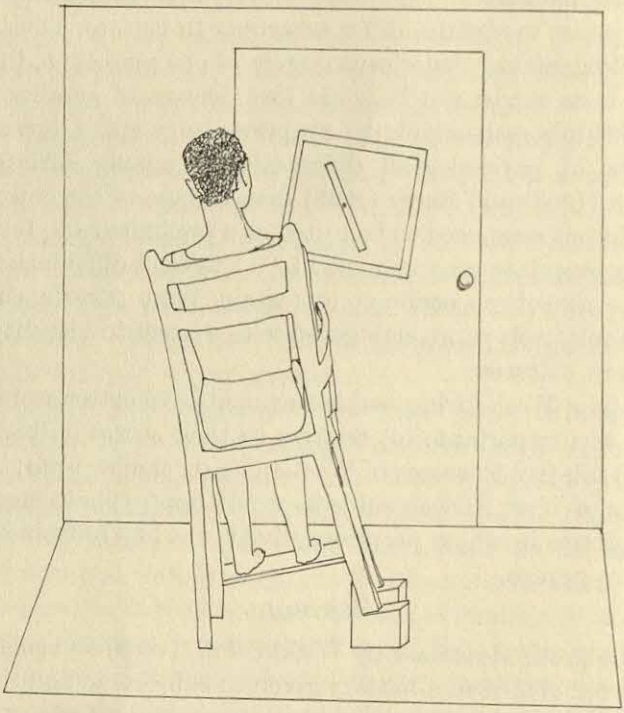


Fig. 1. The Rod-and-Frame test.

to find each item is noted (with a maximum of 5 min allowed) and the total for the test found. The selection of these particular items and their order had been determined by pre-testing among school-children. The procedure in testing was to show each simple figure for 15 sec and then to start timing the period required to find it in the complex figure. If subjects asked to see the simple figure again the complex figure was removed and the simple one shown again for 5 sec, the stop-watch being stopped. No subject asked for more than four repetitions; some asked for one or two, but most often none was requested. The time reckoned was the total search time, without counting time for repetitions of the simple figure. As with the RFT, a high score indicated a poor performance.

The test is clearly one of transactions entirely in the visual field and similar in the skills required for it to Raven's Matrices and the Block of Designs test (used by Dawson, 1963 and Wober, 1966) with which significant correlations were reported.

The subjects

The eighty-six subjects were all male employees of a large company in Southern Nigeria. They were chiefly from Ibo tribal groups, with Edo groups furnishing the others. Both these tribal groups have had similar exposure to Western influences, and both live in geographically similar conditions; both depended for livelihood on subsistence farming and evolved cultures lacking literacy but with intense development of music, dancing and tonal language. All the men were manual workers ranging from completely unskilled to skilled artisans. Literacy ranged from nil

to fair and was influenced by formal education which was assessed by a four point scale. All those with no school experience were marked 1, and those who had only completed 'Primary VI', which was a long established watershed in common educational experience in Nigeria, were marked 3. Those with school experience less than Primary VI were marked 2, and those who had gone on from Primary VI to secondary modern, grammar schools, colleges for teacher training

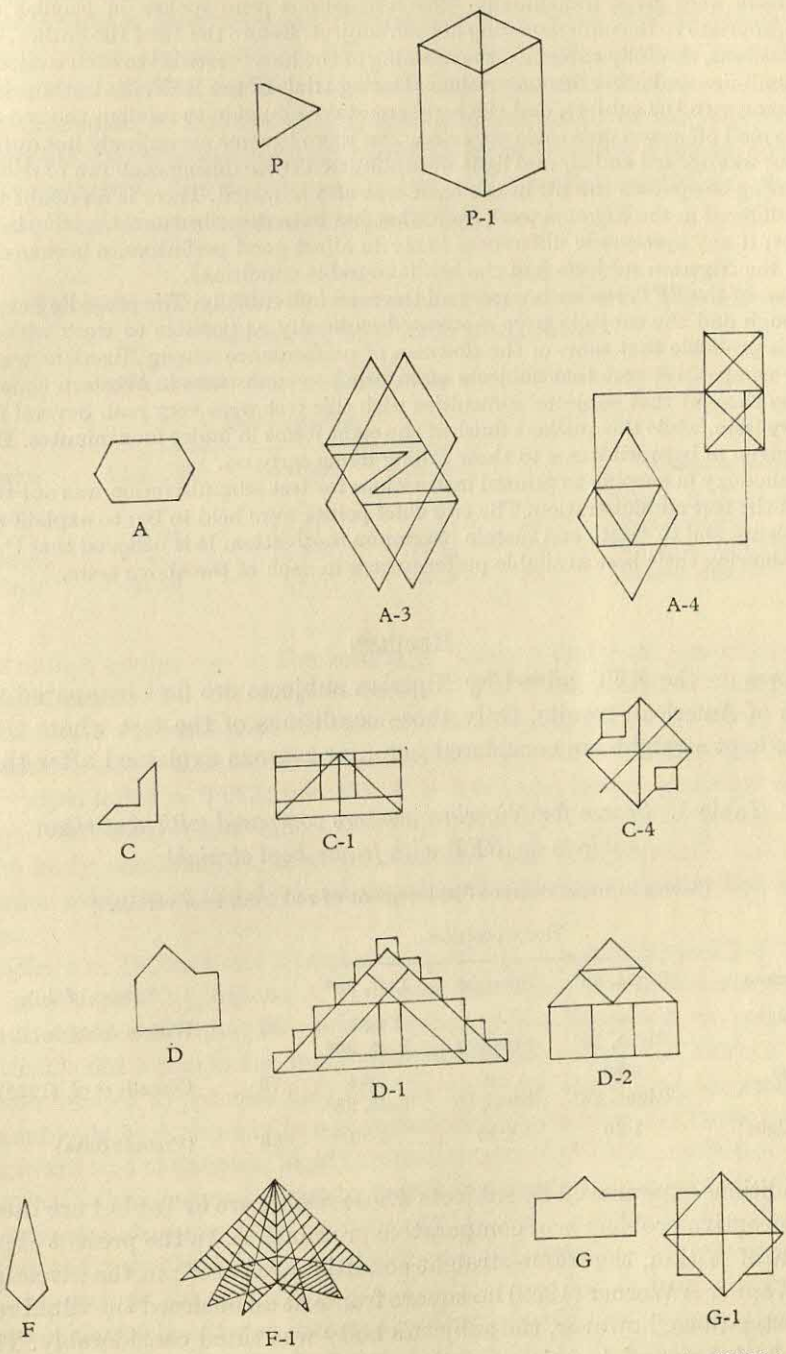


Fig. 2. The selection of items that was used from the Embedded Figures test of Witkin (1950).

or the Company's apprentice school were marked 4. The scale thus did not attempt to represent equal intervals by educational criteria, but the categories were occupied by numbers not greatly skewed in any direction. The average age of the men was 32·2 years.

The testing situations

All the tests were given individually. The instructions were spoken in English or pidgin-English appropriate to the understanding of each subject. Before the RFT the author, aided by a Nigerian assistant, carefully explained the meaning of the idea 'vertical' to each subject with the aid of a plumb-line and other demonstrations. During trials of the RFT, the author remained in the dark room with the subject, and the assistant stayed outside to re-align the rod after each trial and to read off scores on a scale provided. The room became exceedingly hot quite rapidly, and the door was opened and air and light were admitted twice during each run of trials for each subject; during the pauses the tilt of the chair was also adjusted. There is no doubt that social conditions differed in the Nigerian work from what has been described as the setting in American laboratories; if any systematic differences likely to affect good performance occurred, it seems likely that the Nigerian subjects had the less favourable conditions.

In the case of the EFT, the author gave all the tests individually. The preceding explanations were thorough and the subjects were motivated as keenly as possible to work without delay. Though it is probable that some of the slowness of performance among Nigerians was due to a lack of the competitive zest that subjects often bring to such tasks in Western countries, it is nevertheless thought that subjects' difficulties with this test were very real. Several men failed nearly every item, while the quickest finished the eight items in under four minutes. Encouragement was given in between items to those failing items early on.

Rigid orthodoxy in keeping to printed instructions for test administration was not the guiding principle of the test administration. The two chief points were held to be: to explain the nature of the problems, and to create and sustain maximum motivation. It is believed that the subjects were thus showing their best available performances in each of the above tests.

RESULTS

The scores on the RFT gained by Nigerian subjects are first compared with two other sets of American results. Only those conditions of the test where the square frame was kept straight are considered here, for reasons explained after the table.

Table 1. *Scores for Nigerian workers compared with American subjects on RFT with frame kept straight*

(Scores in mean degrees displacement of rod from true vertical.)

Frame	Body position: Right, 28°	Straight	Left, 28°	n	Source of data
Straight	—	—	4·53	53	Witkin & Asch (1948)
	Right, 30°	Straight	Left, 30°		
None	3·30	—	2·64	75	Comalli *et al.* (1959)
	Right, 28°	Straight	Left, 28°		
Straight	1·20	1·18	1·29	86	(Present data)

The conditions experienced by subjects whose results are in Table 1 are those where the proprioceptive problem is of comparative prominence. In the present experiment, and in that of Witkin, the frame-straight condition was used; in the investigation of Comalli, Wapner & Werner (1959) no square frame at all enclosed the tilted rod. In all three investigations, however, the subject's body was tilted considerably. The result is that in both cases of American subjects, the level of error was greater than that

found for Nigerians. Where the body was tilted to the right, the mean Nigerian score was significantly better than for Comalli's subjects ($t = 7{\cdot}60$; D.F. 159; $P < 0{\cdot}01$). Where the body was tilted left, the mean Nigerian score was again significantly better than for Comalli's subjects ($t = 4{\cdot}65$; D.F. 159; $P < 0{\cdot}01$); and the same superiority held good over Witkin's subjects ($t = 3{\cdot}54$; D.F. 137; $P < 0{\cdot}01$).

The conclusion is clear that, in circumstances of what Witkin terms 'analytic functioning' where proprioceptive information (body tilt) is more prominent than the visual inconsistencies, the Nigerians were better than the Americans. One possible interpretation could be that the testing situation was easier for the Nigerians. It has been suggested that this was unlikely owing to social disparity between subjects and English investigator, and to physically hot and stressful conditions; the data of Tables 2 and 3 also argue against this possibility.

Table 2. *Scores for Nigerian workers compared with American subjects on RFT with frame tilted*

(Scores in mean degrees displacement of rod from true vertical.)

	Body position				
Frame	Right, 28°	Straight	Left, 28°	*n*	Source of data
28° Left	—	6·2	9·4	53	Witkin & Asch (1948)
28° Left	11·08	10·86	11·47	86	(Present data)
28° Right	—	5·7	11·98	53	Witkin & Asch (1948)
28° Right	11·16	11·8	12·06	86	(Present data)

The situation giving rise to the results in Table 2 had both proprioceptive and visual displacements for the subjects to deal with. Comparison of Witkin's results for his American subjects with the performance of Nigerian subjects in the body-tilted condition showed no significant differences (for the condition with both frame and body tilted left, $t = 0{\cdot}072$; D.F. 137, $P > 0{\cdot}05$; and for the condition with frame tilted right and body tilted left, $t = 0{\cdot}04$; D.F. 137, $P > 0{\cdot}05$). In the condition where the body was straight, however, and the task was primarily one of having adjustments made to a tilted visual display, the Americans made less error than Nigerians.

The findings in Tables 1 and 2 tend to support hypotheses 2(*a*) and 2(*b*) and therefore, by implication, hypothesis 1. It is also damaging to the argument that the testing situations were unequal and in favour of the Nigerians, to find that they systematically did worse in the body-straight situation shown in Table 2.

A further test of hypothesis 2(*b*) may be made by comparing the results of the Nigerian subjects with those of two groups studied in America by Gruen (1955). One of these groups was of dancers, highly trained in proprioceptive sensitivity; the other was of normal subjects (non-dancers). Gruen classified her results slightly differently, and the present data have been rearranged in Table 3 to parallel her classification.

Gruen stated that there were no significant differences in performance between normals and trained dancers. As in the comparison with Witkin's data, Nigerians were better when body-tilt was involved (proprioceptivity becoming important) and worse than Americans when the body was straight. (Gruen's data were not presented

in a form in which significance of differences could be readily estimated). This is additional support for hypothesis 2 (*b*).

Among the Nigerian subjects only, comparison of performance in conditions where the frame was straight and body tilted, with performance when the frame was tilted with the body kept straight, show that the former elicited much more accurate judgments, i.e. it was easier for them to resolve a proprioceptive than a visual displacement from normal conditions. This supports hypothesis 2 (*a*).

Table 3. *Comparison of RFT scores between Nigerians, Americans and American professional dancers*

(Scores in mean degrees displacement from true vertical.)

Group	*n*	Body and frame tilted same way, 28°	Body and frame tilted opposite, 28°	Body straight frame tilted, 28°	Source
Normals	46	12·33	14·14	7·07	Gruen (1955)
Dancers	30	13·57	15·00	6·58	Gruen (1955)
Nigerians	86	11·32	11·57	11·02	(Present data)

Table 4. *Product-moment correlations between RFT and EFT results for Gruen's American subjects* (*Gruen*, 1955)

	Conditions of RFT		
Group	Body and frame tilted, same way	Body and frame tilted, opposite ways	Body straight frame tilted
Normals	0·47**	0·43**	0·75**
Dancers	0·32	0·32	0·36*

* $P < 0{\cdot}05$ for difference from zero. ** $P < 0{\cdot}01$.

Table 5. *Product-moment correlations between RFT and EFT and educational level for Nigerian subjects* ($n = 86$)

Variate	RFT	Educational level
EFT	0·21	−0·37*
RFT	—	0·08

* $P < 0{\cdot}05$ for difference from zero.

Gruen gave her two groups of subjects the EFT as well as the RFT, and the correlations between the tests are available. The Nigerians also did both tests, and so the patterns of inter-correlations can be compared.

The evidence from these correlations can be interpreted as support for hypothesis 3. Thus in Gruen's investigations, for those parts of the RFT involving body-tilt (proprioceptive problems) the dancers showed no significant correlation with their performance on the EFT. That is to say, the two tests posed to them problems which were dealt with in basically different ways. In the condition where the body was kept straight (the problem principally visual) there was a significant correlation for dancers with their EFT results. In any case, the situation for normals was different from that for dancers. The Nigerian subjects showed no significant correlation between EFT

and RFT results, thus showing their closer similarity to American professional dancers than to other American subjects. This appears to be good support for hypothesis 3. It is also shown among Nigerian subjects that formal (Western) education has more bearing upon EFT that upon RFT results.

Discussion

It has been argued that Witkin's Rod and Frame test contains a proprioceptive component which makes it essentially different from tests involving purely visually mediated transactions. This being so, groups of subjects who might by culture or by training be especially sensitive to the proprioceptive field would show characteristically different patterns of scoring on this test. This has been found to be the case. American dancers, and especially Nigerian manual workers, have shown that for them the RFT does not measure the same skills or analytic style as the EFT.

This finding has two areas of applicability. The first is that there exists a test of 'analytic functioning' (Witkin's term, 1962) in which Africans have been shown to do as well as, or better than Americans or Europeans. This is a rare finding that has occurred once only before in the work of Beveridge (1939) who compared fifty-two Ghanaian students with twenty Europeans on a test of adjusting a rod to the horizontal, inside a tilted room. Though half the Europeans were women and their average age nearly twice that of the Africans, rendering the results equivocal, nevertheless in this task involving action in a proprioceptive field the Africans did better than the Europeans. Beveridge wrote (op. cit.) 'their perception is guided less by visual and more by other cues than is that of the European.'

The second application of the present findings is the modification of Witkin's theory of psychological differentiation. It would appear that 'style of cognitive functioning' is not so uniform throughout all fields of an individual's expression as had originally been supposed by Witkin. The finding in America that the EFT and similar visual tests indicated a person's level of psychological differentiation was supported in Sierra Leone by Dawson (1963) using visual tests. However, visual tests appear not to be sole indicators of psychological differentiation. The evidence here is that such differentiation may occur in sensory fields other than the visual one. Using the idea of sensotypes introduced by Wober (1966) it may be suggested that suitable tests in visual idiom indicate levels of differentiation implying certain sorts of socialization experience and social behaviour. Tests in proprioceptive idiom, however, may be more appropriate for measuring differentiation among those of a less visual sensotype and may be associated with experiences of socialization and with social behaviour in patterns that have not as yet been worked out.

References

Bennett, D. H. (1956). Perception of the upright in relation to body image. *J. ment. Sci.* **102**, 487–506.

Berry, J. W. (1965). A study of Temne and Eskimo visual perception. Preliminary Report no. 28, University of Edinburgh Department of Psychology.

Beveridge, W. M. (1939). Some racial differences in perception. *Br. J. Psychol.* **30**, 57–64.

Comalli, P. E., Wapner, S. & Werner, H. (1959). Perception of verticality in middle and old age. *J. Psychol.* **47**, 259–66.

DAWSON, J. L. M. (1963). Psychological effects of social change in a West African community. Unpublished MS., Department of Social Anthropology, University of Edinburgh.

EPSTEIN, L. (1957). The Relationship of Certain Aspects of Body Image to the Perception of the Upright. Unpublished Doctors Dissertation, New York University.

GRUEN, A. (1955). Dancing experience and personality in relation to perception. *Psychol. Monogr.* **69**, no. 14, whole no. 399.

PRESTON, C. E. (1964). Psychological testing with Northwest Coast Alaskan Eskimos. *Genet. Psychol. Monogr.* **69**, 323–420.

RIESMAN, D., GLAZER, N. & DENNEY, R. (1953). *The Lonely Crowd: A Study of the Changing American Character.* New York: Doubleday.

WITKIN, H. A. & ASCH, S. E. (1948). Studies in space orientation, IV: further experiments on perception of the upright with displaced visual fields. *J. exp. Psychol.* **38**, 762–82.

WITKIN, H. A. (1950). Individual differences in ease of perception of embedded figures. *J. Personality* **19**, 1–15.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., BRETNALL-MEISSNER, P. & WAPNER, S. (1954). *Personality Through Perception.* New York: Harper.

WITKIN, H. A., DYK, R. B., FATERSON, H. E., GOODENOUGH, D. R. & KARP, S. A. (1962). *Psychological Differentiation.* New York: Wiley.

WOBER, M. (1966). Sensotypes. *J. soc. Psychol.* **70**, 181–9.

YOUNG, A. H. (1959). A test of Witkin's field dependence hypothesis. *J. abnorm. soc. Psychol.* **59**, 188–92.

(*Manuscript received* 15 *April* 1966; *revised manuscript received* 24 *November* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 39–49
Printed in Great Britain

SOME EXPERIMENTS RELATING TO BARTLETT'S THEORY OF REMEMBERING

By ALAN GAULD
Department of Psychology, University of Nottingham

and GEOFFREY M. STEPHENSON*
Department of Psychology, University of Keele

Bartlett's theory that remembering is a 'reconstructive' process is based largely upon the ways in which subjects change and distort prose passages when reproducing them from memory. If such changes and distortions are to serve as the foundation for a theory of remembering it is clearly desirable to be quite certain that the persons who make them really are trying to remember, and are not deliberately inventing material to fill in gaps in their memories. Three experiments were carried out in an attempt to shed light on this question. In the first experiment it was ascertained that strict instructions to avoid errors markedly lessened the number of errors which subjects make in reproducing prose passages. In the second it was found that subjects are fairly adept at picking out their own errors. In the third it was shown that conscientious subjects make notably fewer errors than less conscientious ones. It is concluded that for various reasons subjects in Bartlett-type experiments often do more or less consciously invent material to bridge gaps in their memories, and that Bartlett's claims that the occurrence and nature of such changes and distortions shed light on the mechanisms of remembering must be regarded with caution.

Bartlett's experiments on remembering (Bartlett, 1932) are too well known to need a detailed exposition. The most celebrated are probably those on the 'repeated reproduction' and the 'serial reproduction' of prose passages. In the former each subject was given a prose passage to read, and was then required to reproduce it in writing a number of times; first, after an interval of about 15 min, and subsequently as opportunity offered. In the latter a subject would be given a prose passage to read and would be asked to reproduce it after a short interval; his reproduction would be given to a second subject to read and reproduce; the second subject's reproduction to a third; and so on. The results of these two kinds of experiment may be summarized briefly. (*a*) Subjects, or at least subjects from our society, tend to abbreviate prose passages considerably, especially by the omission of material which smacks of the odd or the incongruous, (*b*) Subjects are likely to reduce prose passages to a conventional form, not merely by the omission of irregularities, but by changing and distorting material, and even by inserting new material of their own. Incongruities, together with ideas, facts or sentiments alien to our culture, may thus be smoothed over or rationalized away; abstract or difficult notions or arguments may be transformed into simple and concrete ones; and stories which have no point or purpose for a British audience may be injected with clichés and rounded off with an obvious moral. (*c*) The general setting of a passage is remembered, especially its 'affective tone', and so are certain dominant details; but in between the dominant details the processes of rationalization and conventionalization have free play, and departures from the original are common, often in directions suggested by its 'affective tone'.

On the basis of differences between originals and reproductions which these experi-

* Now at Department of Psychology, University of Nottingham.

ments revealed, Bartlett erected a theory of remembering which has since been influential (on the place of the theory in the history of recent psychology see Oldfield & Zangwill, 1942; on the genesis of the theory see Bartlett, 1958, ch. 8). The general argument is that rote memory, if it occurred frequently, would be biologically wasteful; but in fact the experimental results show that it rarely occurs. The material which emerges when one tries to remember a particular event or set of events is always altered by one's experiences of other, related events. Bartlett (1932) therefore postulated that memory is not 'primarily or literally reduplicative, or reproductive' (p. 204). Only a few outstanding facts, the 'dominant details', are straightforwardly remembered. Normally an incoming event is assimilated to the 'schema' formed by similar incoming events in the past.

Schemata seem to be poolings or averagings of such incoming events, reminiscent of 'generic images' or composite photographs, but conceived on the more active lines of Head's body-scheme. If, however, memory is usually schematized in this way, two problems arise: (*a*) how is it that reproductions of prose passages are full of highly specific facts; and (*b*) how are specific facts recalled in their correct temporal order?

Bartlett's answer was that 'an organism has somehow to acquire the capacity to turn round upon its own "schemata" and to construct them afresh' (p. 206). Given that certain dominant details are remembered (seemingly as images) missing items, and the order of missing items can be inferred or reconstructed from generalized schemata; (though, of course, the inferences will rarely be perfect, and changes of the kind shown up in reproductions of prose passages are bound to occur). 'The need to remember becomes active, an attitude is set up; in the form of sensory images, or, just as often, of isolated words, some part of the event which has to be remembered recurs, and the event is then reconstructed on the basis of the relation of this specific bit of material to the general mass of relevant past experience or reactions, the latter functioning...as an active organised setting' (p. 209).

Rationale of experiments

It appears that in his experiments on the repeated reproduction and the serial reproduction of prose passages Bartlett (1932) did not give any very specific instructions to his subjects: 'I thought it best, for the purposes of these experiments, to try to influence the subjects' procedure as little as possible' (p. 78); and on p. 175 he hints that the mere fact of being in an experimental situation will provide subjects with an incentive to strive after accuracy. Most people who retell a story are unlikely to care very much whether the story which they retell is the same, detail by detail, as the story which they originally heard. In other words, they are most unlikely to take pains that what they come out with is always what they remember rather than what they guess at or even consciously invent. Now if the changes and inventions in reproductions of stories (and of other prose material) are to serve as the foundation for a theory of remembering, it is clearly desirable to establish that when the changes and inventions were made those who made them were indeed seriously trying to remember, and were not more or less consciously guessing or romancing in order to fill in gaps in their memories. The problem is to distinguish between these two possibilities. Three ways in which this might be done suggest themselves. (*a*) A

group of subjects might be given specially strict instructions not to bridge gaps in their memories by guesswork, and the number of their errors compared with that of a control group given no special instructions. (*b*) Subjects might be urged to pinpoint their own deviations from the original. (*c*) An investigation might be conducted to determine whether highly conscientious persons, who would do their very best to be accurate, make fewer errors than less conscientious persons, who might be expected to invent what they cannot remember. These considerations led to the following experiments.

Experiment I

The first experiment was a pilot investigation. It was carried out to see whether subjects were able to pick out their own errors in a passage they had reproduced, and whether instructions of varying degrees of strictness affected the accuracy of the reproductions.

Method

Subjects. The subjects were seventy-five students, in three groups of twenty-five each.

Procedure. The passage chosen was *The War of the Ghosts* (Bartlett, 1932), and the method of serial reproduction was used. The experimenter read each successive version of the story to subjects individually. Instructions were given to subjects immediately after the reading of the story, and different instructions were used for each of the three groups.

Group I. 'Write down the story I've just read you as exactly as you can.'

Group II. 'I want you to set down on paper as much as you can remember about the story I have just read you. It is particularly important that you don't set down facts and incidents which were not in the original. If there are parts of the original that you just do not remember, do not invent material to fill the gaps for the sake of being able to tell a complete story; leave blanks. If there are facts and incidents which you suspect were mentioned in it, but about which you are not quite certain, set them down in brackets.'

Group III. 'I want you to reproduce this story as best you can in the following way. Take great care that each item you set down corresponds with something in the story I read you—not word for word, of course, but fact for fact. So that if you come to something you can't remember don't fill in to make a story of it; leave a blank. And, if you come to something you're not quite certain about, put it in square brackets. I want you to look on this not as a test of memory, but as an exercise in being as scrupulous and as honest as you can in deciding what is and what isn't in the original story. If you have any doubt over any item, come down on the side of doubt, don't come down on the side of certainty.'

When each subject had completed his version of the story he was given the following additional instructions: 'Now I'm going to run through your version of the story incident by incident and fact by fact, and ask whether or not you're *quite certain* that all the facts and incidents you set down were really in the original story. Do not say that you are certain so long as you entertain the smallest doubt.'

The experimenter then read the subject's version in as neutral a tone as voice of possible, pausing for a second or two between phrases to allow him time to have second thoughts if he wished. Phrases about which he expressed doubt were enclosed in square brackets.

Each series of reproductions was arbitrarily terminated when a form of story was reached which seemed to the experimenter to offer little scope for change. No subject was required to reproduce a story of less than 100 words in length, and the average length of penultimate stories was 135 words. There were five series each in groups I and II and seven series in Group III.

Measurement of performance

To evaluate differences between the three groups, an 'error rate' was calculated for each subject. This was defined as the average number of errors for one hundred words of reproduction. An error was defined as any deviation whatsoever from the version of the story which a subject

heard, with the following qualifications: (*a*) synonyms and synonymous phrases were accepted; (*b*) omissions were never penalized; (*c*) general words were accepted in place of particular ones, e.g. 'boat' for 'canoe'; (*d*) time order errors were ignored, being too difficult to score; (*e*) obvious mishearings, e.g. 'eels' for 'seals' were never penalized; (*f*) place name errors were never penalized; (*g*) if an error recurred, it was penalized only once.

Otherwise, all inaccuracies were penalized, including such plausible insertions as 'savage' in the description of the fight, and inversions like 'up river' for 'down river'. Hence, the error rate was not a measure of the overall efficiency of recall, but an indication of the extent to which a reproduction, however lengthy or adequate, contained material not in the original, i.e. material of the kind upon which Bartlett based his theory of remembering. A 'revised error rate' was obtained from each subject's corrected version. This was calculated on the basis of the number of errors and words remaining outside the brackets in which he had enclosed 'erroneous' material. Thus it was possible for error rate to increase if correct material was enclosed in brackets.

Kay (1955, pp. 83–4) has pointed out the difficulties of quantifying meaningful material if the figures obtained are not to mislead; the natural quantitative units, words and sentences may not adequately express the quality of a reproduction. He made, for his passages, both a literal verbal analysis and a content analyis in which the exact fidelity of vocabulary was not important. 'Errors' featured only in his verbal analysis and were defined as 'words substituted in the reproduction for those in the original' regardless of their meaning. Such a measure makes sense only if subjects have been asked specifically to reproduce the story in the exact words of the original. Strictly verbal deviations from the original were not our concern. We were concerned with falsehood, and the 'error rate' was designed to indicate the proportion of false or untrue statements in the reproductions as a whole. The method, however, still has the disadvantage of giving equal weight to all errors. It is hard to see how this could be overcome. To illustrate the scoring method a reproduction of the original *War of the Ghosts* story is printed in an appendix (p. 49) with the errors printed in italics (for the original passage, see Bartlett, 1932, p. 65).

Table 1.

(A) *Intergroup differences in initial error rate*

Group	Number of series	Total number of words	Total number of errors	Individual error rate	
				Mean	S.D.
I	5	4451	208	4·68	2·14
II	5	3746	130	3·29	1·74
III	7	4422	122	2·78	1·77

(B) *Analysis of variance on differences in error rate*

Source	D.F.	Sums of squares	Variance	S.D.
Between	2	48·41	24·20	—
Within	72	268·50	3·59	1·89

RESULTS

Table 1A gives details of initial error rate. Analysis of variance (Table 1B) shows the overall difference to be significant ($F = 6{\cdot}74$; $P < 0{\cdot}005$). The differences between groups I and II and between groups I and III are significant ($t = 2{\cdot}57$; DF $= 49$; $P < 0{\cdot}02$; and $t = 3{\cdot}5$; DF $= 49$; $P < 0{\cdot}01$, respectively). The difference between groups II and III does not reach a significant level ($t = 0{\cdot}94$; DF $= 49$).

In Table 2A differences in revised error rate are presented. The overall difference between the groups remains significant ($F = 5{\cdot}45$; $P < 0{\cdot}01$), as may be seen from Table 2B. Table 2C shows the differences between initial and revised error rate separately for each of the three groups. All differences are highly significant, using *t* tests for correlated means.

Table 2.

(A) *Intergroup differences in revised error rate*

Group	Total number of words	Total number of errors	Individual revised error rate	
			Mean	S.D.
I	4003	114	2·93	1·72
II	3303	64	2·00	0·98
III	3828	61	1·62	1·44

(B) *Analysis of variance on revised error rate differences*

Source	D.F.	Sum of squares	Variance	S.D.
Between	2	22·67	11·34	—
Within	72	149·49	2·08	1·44

(C) *Intergroup differences between initial and revised error rates*

Group	Mean difference	S.D.	D.F.	t	P
I	1·75	1·11	24	7·63	< 0·001
II	1·29	1·20	24	5·26	< 0·001
III	1·16	1·16	24	4·92	< 0·001

Discussion

It is clear from the results that strict instructions to remember reduce the error rate considerably, and that the opportunity to express doubts about parts of reproductions reduces it still further. It is possible that in the second part of the experiment subjects were able to pick up cues from the experimenter's tone of voice when he read back their version of the story. In the first part of the experiment the method of serial reproduction may have introduced some distortion into the results. There were seven series of reproductions in group III, but only five in groups I and II. Both sets of results seemed to be clear in their implications, but they needed verification under more strictly controlled conditions.

Experiment II

Strictness of instruclions and accuracy of recall

Method

Subjects. The forty-seven subjects in this experiment were the members of four psychology practical groups, organized alphabetically.

Procedure. Four sets of instructions were devised and assigned randomly to the four groups. Because many of the subjects were familiar with Bartlett's story, an unfamiliar Victorian 'true ghost story' was used (copies may be obtained from the authors on request). This was prerecorded, together with all the instructions. Subjects were told: 'Listen carefully to the following story. You will be asked later on to reproduce it'. They then heard the story, after which the following instructions were given. (The point at which the recorded instructions finished for each group is indicated in parentheses):

'Now write out the story I have just told you' (group I) 'accurately' (group II). 'You may, of course, use your own words or phrases, so long as they correspond to those in the original. Give as full an account as you can' (group III). 'Make every possible effort to be accurate, and

take special care not to introduce words or phrases to which nothing in the original corresponds. Please follow these instructions scrupulously' (group IV).

Measurement of performance

It was planned to make a comparison of errors between groups I and II, using the 'error rate' as defined previously, and again between groups III and IV. Instructions for group II differed from instructions for group I only in the added injunction to be 'accurate'. Instructions for group IV differed from those for group III only in the emphasis placed on obeying the instructions. It was expected that group II would have a significantly lower error rate than group I, and group IV a significantly lower error rate than group III.

Table 3. *Significance of error rate differences between groups I and II, and groups III and IV using* t *test for independent means*

	n	Mean	S.D.	t	D.F.	P
Group I	13	4·184	1·597	2·513	23	< 0·02
Group II	11	2·651	1·189			
Group III	11	3·476	1·757	0·950	22	N.S.
Group IV	12	3·074	1·192			

Results

The predicted differences in error rate were tested for significance (see Table 3). As the results show, the hypothesis was supported by the difference between groups I and II. The difference between groups III and IV was not significant. The addition of the single word 'accurately' to the simple instructions to write out the story, appears to have had a striking effect. Presumably, the subjects in group I felt no obligation to be accurate, being concerned more to produce a convincing, if fanciful, story. Subjects in groups III and IV were all enjoined to be accurate and were told they could use their own words so long as they corresponded to those in the original. The additional request to obey the instructions in group IV appears to have had no effect over and above the initial injunction to be accurate.

EXPERIMENT III

The awareness of inaccuracy in recall

Method

Subjects

The subjects were seventy-six undergraduates following an introductory psychology course in their first year.

Procedure

Bartlett's story, *The War of the Ghosts*, was used. The story was read by the experimenter, after which the subjects were asked to answer ten prepared written questions on the story: (1) When did the two young men from Egulac set out? (2) Where did they go before they encountered the canoes? (3) What was the object of their excursion? (4) What change occurred in the weather before the young men heard war cries? (5) Where did they hide when they heard war cries? (6) How many canoes came up to the two young men? (7) Did the young man who said he had no arrows eventually go with the men? (8) What made the young man think his companions were ghosts? (9) What was the first thing the young man did when he returned to his home? (10) What did the young man do immediately after telling his story?

The written instructions were: 'Answer each question in the space provided below it. Please answer ALL the questions—do not leave blanks or write "don't know". When you have answered each question please indicate by putting a tick in the appropriate column at the right whether you are *quite certain*, *fairly certain* or *doubtful* that your answer is correct, or whether you are just *guessing*.'

The subjects then expressed their degree of certainty about the correctness of each answer. It was expected that there would be a very strong association between subjective estimates of certainty, and actual correctness of answers.

Table 4. *Ratings of accuracy and actual correctness of answers*

Answers	Quite certain	Fairly certain	Doubtful	Guessing	Totals
Correct	262	109	54	11	436
Incorrect	65	89	97	83	334
Totals	327	198	151	94	770
$(O-E)^2/E =$	73·65	0·18	26·69	77·11	$\chi^2 = 177·6$

Results

Table 4 shows the number of correct and incorrect answers in the four possible accuracy ratings. All seventy-seven subjects answered all questions and rated them accordingly. The total $\chi^2 = 177·6$ shows the association between accuracy rating and correctness of answers to be reliable at a very high level ($P < 0·001$).

When subjects reported that they were 'quite certain' of their answer, their answers were correct 80 % of the time; when 'fairly certain' their answers were rather more likely to be correct than not (55 % accuracy); but when 'doubtful', accuracy fell to 36 %, and when 'guessing only 12 % accuracy was achieved, A biserial correlation coefficient of 0·61 indicates the magnitude of association between the two variables.

Experiment IV

The relation between strength of conscience and errors in recall

The results of the pilot study seemed to indicate that the tendency to make errors, thereby disobeying the instructions, is largely a product of the subject's attitude to the remembering task. Errors could, it seemed, be avoided, if the subject was so inclined. Could the tendency to produce errors be related to some more general factor of conscientiousness? Are errors produced only by the unscrupulous? Stephenson (1963, 1966) had previously used an interview method for the assessment of strength of conscience. Psychopaths scored significantly lower, and in a large group of 15-year-old schoolboys there was a significant association with rule-breaking behaviour.

Methods

Interview procedure

Interview schedule was slightly modified to make it more suitable for use with students. The questions centred on three factors of conscience: first, commitment to altruistic principles of conduct; secondly, the tendency to suffer from guilt feelings and remorse; thirdly, anxiety about the approval of other people.

General procedure

Forty-one first year students in an introductory psychology course were told *The War of the Ghosts* and asked to 'reproduce it as accurately as you can'. They were then interviewed indi-

vidually over a period of two months. The interviewer had no knowledge of their reproductions. Each interview lasted from 20 to 40 min, and was tape-recorded. Assessments were made from the recordings. The subjects then completed the AH 5 intelligence test. Conscientiousness and error-production might both be related to intelligence, and it was thought necessary to control for this. It was predicted that there would be a strong positive relation between error rate and strength of conscience as assessed from the interviews, and that this would be independent of intelligence. In order to elucidate the meaning of 'error rate' in relation to the overall accuracy of the reproduction, an 'accuracy score' was calculated for each reproduction. Fifty main points were culled from the original passage, and a subject's accuracy score was the number of points correctly made in his reproduction.

Table 5. *The relationship between errors in recall, accuracy of recall, conscience strength and intelligence test scores* ($n = 41$)

	Mean	Variance	Product-moment correlations 2	3	4
1. Error rate	4·654	2·676	−0·365	−0·201	−0·400
2. Conscience strength	28·927	48·848	—	−0·049	−0·127
3. AH 5 total score	40·024	71·975	—	—	0·157
4. Accuracy score	28·805	44·742	—	—	—

Results

Table 5 shows the correlations between error rate, conscience strength, intelligence test scores and accuracy scores. The correlation between error-rate and conscience strength is significant ($t = 2{\cdot}45$; D.F. $= 39$; $P < 0{\cdot}01$).

There is a slight tendency for more intelligent subjects to make fewer errors, but this tendency is not significant ($t = 1{\cdot}28$; D.F. $= 40$; $0{\cdot}125 > P > 0{\cdot}10$). Intelligence is virtually unrelated to the measure of conscience strength. This leaves the initially significant negative correlation between conscience strength and error rate at a significant level when intelligence, as a possible factor in the result, is partialled out ($r_{12.3} = -0{\cdot}363$). The prediction is confirmed. The more conscientious the subject the less likelihood is there that he will make errors in recall.

'Accuracy' was significantly and negatively related to 'error rate' but showed no significant relation with either AH 5 scores or with the measure of conscience strength.

DISCUSSION

The absence of correlations between intelligence test scores and error rates, and between intelligence test scores and quantity of material correctly reproduced, is perhaps just part of the general scarcity of correlations between the intelligence test scores of university students, and their performances in other situations.

The inverse correlation between the quantity of material correctly recalled and the number of errors made was only to be expected. But it is not entirely clear that Bartlett would predict it. Those people who are most successful at remembering show fewest signs of the processes of 'construction' which Bartlett claimed to be its essential basis. Do such people make use of 'rote' memory, which Bartlett (1932, pp. 204–5) thinks wasteful and biologically primitive? Or are they unusually skilful at the task of reconstruction? The former hypothesis will hardly stand up; people with good memories are not generally primitive or backward in any way. The latter hypothesis would seem to be quite untestable. The fact to be explained is that people

with capacious memories do not on the whole make many errors. Yet it is upon the occurrence of errors that the theory that remembering is a reconstructive process is principally based.

The striking ability to pick out their own errors which subjects display, the distinct inverse correlation between error rate and conscientiousness, and the fact that strict instructions to be accurate decidedly reduce the error rate, all combine to suggest rather strongly that the ordinary subject in experiments of the Bartlett kind does not do his best to *remember*. On the contrary, he probably bridges gaps in his memory with a great deal of more or less conscious guessing and inventions. There is, of course, bound to be an intimate relation between processes of guessing and invention, and processes of remembering; but to draw, as Bartlett has, conclusions about the nature of remembering from those parts of reproductions of prose passages which are most likely to be conscious guesses or inventions seems somewhat incautious. If an outstandingly conscientious person, whose errors would certainly not be the result of conscious guessing or invention, distorted prose passages despite great moral pressure to be accurate, then indeed one could start enquiring whether some basic feature of memory mechanisms were being laid bare. So far no experiments have specifically met these conditions. From the results presented above and from the results of Dallenbach (1913, listed in Hunter, 1964, p. 175), it seems probable that even highly conscientious persons under great pressure to be accurate would still make a few errors; but if the errors were only few (and they would, we feel, certainly be many fewer than in the majority of the reproductions quoted by Bartlett; Kay, 1955, p. 95, has emphasized the substantial accuracy of subjects' reproductions), they would hardly constitute a sound basis for a theory about the constructive nature of remembering. One would be tempted to regard them as 'illusions of memory', as special cases to be explained away, rather than as clues to fundamental features of remembering.

The experiments described in this paper might be thought to cast doubt on Bartlett's theory in another way. Subjects in experiments on remembering seem to be quite successful at picking out from their reproductions parts which are demonstrably guesses, inventions or inferences. Since, generally speaking, only a small proportion of the material is picked out in this way, it might reasonably be concluded that the bulk of any given reproduction is unlikely to be the product of guesswork or inference. The hypothesis that it *is* guesswork or inference is at any rate unverifiable; for the largest part of the average reproduction is neither demonstrably out of accord with its original, not yet alleged by its author to be the result of guesswork.

It might be replied that when subjects pick out faulty parts of their reproductions they are simply pinpointing the unjustified guesses and the weak inferences. The parts which are not eliminated are, for the most part, inferences so commonplace and so obviously sound that no doubts about them are likely to arise. To this the obvious answer is that the parts of reproductions which subjects pick out as doubtful (and which are in fact wrong) generally constitute inferences of the soundest and most plausible kind. The parts of reproductions which subjects do not pick out as doubtful (and which are in fact correct) very frequently contain statements which, regarded as guesses or inferences, are highly implausible.

CONCLUSION

It will be clear from the discussion of results that we feel that our experiments to some extent undermine Bartlett's theory of the reconstructive nature of recall. An alternative way of looking at the whole problem of the errors in reproductions of prose passages would be this. People who are required (say) to retell a story are under strong pressure to produce something completed and coherent, a chain of particular events rather than the series of vague statements ('it was a story about Indians') which they would perhaps more readily vouch for. Their propensity to invent and distort material to this end may be lessened or restrained by a number of factors. One is possession of a good memory, where by memory is meant the ability to recall (without processes of 'construction') both general and particular aspects of past experiences. A person with a good memory does not have to invent or distort material. Another is the nature of the situation; whether an experiment is going on, whether strict instructions have been given, and so forth. Yet another is, of course, conscientiousness. If a subject does not have a good memory, is not under moral pressure, and is not conscientious, the urge to tell a story, and to be particular rather than general, will have its way, and he will probably distort the original passage markedly. Furthermore, he will not trouble to enquire very closely whether or not he has made any errors. He may be quite satisfied with his efforts, so long as he can feel that the reproduction is in a general way pretty much like the original.

Bartlett may well have been entirely right in supposing that what might be called 'generic' memory mechanisms are more fundamental than mechanisms for the recording of particular events. The above objections are only to his view that such 'generic' memory mechanisms are put to work in the 'reconstruction' of chains of particular past events.

Finally, it appears to us that Bartlett's approach to the explanation of human behaviour is the right one in an important general respect. His theory of remembering has often been criticized on the grounds that to talk of an organism 'turning round on its schemata' and 'constructing them afresh' implies that there must be a ghostly homunculus inside its head to carry out the said process of reconstruction (cf. Oldfield & Zangwill, 1942, pp. 122–3). Now, of course, explanations in terms of homunculi are merely regressive; but at the same time Bartlett's standpoint was an enormous advance upon that of Ebbinghaus and traditional associationists. Associationists have always tended to portray human beings as subjects who undergo or live through experiences; to regard them in short as patients. Bartlett thought of human beings as active, organizing, striving, intending agents. He was surely right about this; and he may well be right in much the same way as against the neo-behaviourists and machine theorists of today. Oldfield (1954) and Broadbent (1958, pp. 61–7) have both given machine-theoretical accounts of Bartlett's views on schemata; both have concentrated on his notion that schematization economizes on storage space, and have said little of the dynamic side of his theory. It may be that if we are not to have agents in our machines, we must find some way of making agents out of our machines. The lessons which Bartlett taught have not yet been perfectly assimilated.

We are indebted to Dr R. B. Joynson and Dr L. J. Newson for their helpful comments and advice.

Appendix

Scored Reproduction of 'The War of the Ghosts'

'The War of the Ghosts'

'Two young men from Egulac went down to the river, at night, it *was* foggy and cold, the water was *muddy*. Then *two* Indians came *across the river* in a canoe and said 'We are going down the river to make war, would you like to come along'. One of the young men said he was unable to go because his relatives did not know where he was and he did not *want* to be killed. But he turned to his companion and said 'but you may go *my friend*'. His friend said he had no arrows. The Indians told him that there were arrows in the canoe. So he *climbed into the canoe* and sailed up the river to a town on the other side of Capabla. The people of this town came down to the river and began to fight. Many people were killed. He was told he was hit, realised they were ghosts, then *saw* Indians *climbing back* into boats so *went back* with them. He *says* he did not feel sick. *Walked* back up *the hill* from the river to his house and told *his family* what had happened. He said he had been fighting with the ghosts, had been hit but was not sick. He *went to bed* but next morning he suddenly *felt peculiar*, fell down, something black *rolled* out of his mouth and he was dead.'

Number of errors = 19; number of words = 238; error rate = 7·98.

References

Bartlett, F. C. (1932). *Remembering*. Cambridge University Press.

Bartlett, F. C. (1958). *Thinking*. London: Allen and Unwin.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

Hunter, I. M. L. (1964). *Memory*. Harmondsworth: Pelican Books.

Kay, H. (1955). Learning and Retaining Verbal Material. *Br. J. Psychol.* **46**, 81–100.

Oldfield, R. C. (1954). Memory mechanisms and the theory of schemata. *Br. J. Psychol.* **45**, 14–23.

Oldfield, R. C. & Zangwill, O. L. (1942). Head's concept of the schema and its application in contemporary British psychology. Part III. Bartlett's theory of memory. *Br. J. Psychol.* **33**, 113–29.

Stephenson, G. M. (1963). Conscience and Psychopathy. Ph.D. Thesis, University of Nottingham Library.

Stephenson, G. M. (1966). *The Development of Conscience*. London: Routledge and Kegan Paul.

(*Manuscript received* 16 *November* 1966)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 51–56
Printed in Great Britain

AFFECTIVE CONNOTATION OF WORDS AS A FACTOR IN FREE RECALL LEARNING

By JOHN A. McNULTY and CATHERINE J. E. ISNOR*
Dalhousie University, Halifax, Canada

Studies concerned with the relation between affective connotation and memory have usually shown that pleasant words or experiences are recalled better than unpleasant or neutral ones. Differences are often small, however, and exceptions are common, perhaps because of the confounding effects of other variables such as familiarity, frequency, word length and context; these variables are known to affect learning, but have frequently been uncontrolled in these studies. The present experiment was an attempt to determine how affective connotation of the material influences free recall learning when the possible confounding effects of some of these other variables are eliminated. Students from a large class rated 120 words on a seven-point pleasantness/unpleasantness scale. From these ratings, a pleasant, unpleasant, and neutral list of words were constructed and given to other subjects to learn. Rated familiarity of the words, frequency, word length, and context were also controlled. Results of the study showed that pleasant and neutral words were learned more easily than unpleasant ones, but, contrary to the results of some studies, there were no differences between pleasant and neutral words.

Interest in the relation between affective connotation and memory has a relatively long history in psychology. The emphasis in early studies (reviewed by Meltzer, 1930; Cason, 1932; Barrett, 1938) was on personal experience, and, although results were far from conclusive, most of them showed that pleasant words or events were recalled better than unpleasant or neutral ones. Later studies have employed more objective procedures, including having subjects rate for 'pleasantness' the materials which they were going to learn. Results of these studies have again tended to be inconclusive.

In general, affective connotation of the learning material appears to have little or no overall effect upon paired-associate learning (e.g. Balken, 1933; Ratliff, 1938; Keppel, 1963). Because several different sub-processes are involved in paired-associate learning (see, for example, Underwood, Ekstrand & Keppel, 1965), it is possible that the lack of any overall effect might be the result of a 'cancelling-out' of facilitatory and inhibitory effects operating on the various sub-processes. At this stage, it is impossible even to speculate how affective connotation might affect these sub-processes. The primary concern of this paper, therefore, is with the influence of affective connotation upon the single-process learning situation of free recall learning in which subjects are simply given a number of verbal units to learn for a number of trials.

Some free recall studies have shown that pleasant items are recalled better than neutral or unpleasant ones, and some have shown little difference. One factor which may account for part of the inconclusiveness in the results of experiments to date is 'context'. Typically, more than one type of material is included in the learning list which subjects are given to learn (e.g. Stagner, 1933; Amster, 1964). As McGeoch & Irion (1952, p. 386) have pointed out, when 'pleasant' and 'unpleasant' or 'neutral' words are combined in the same learning list, the effect that the affective connotation of the words will have in any given experiment could depend upon the particular

* Now at the Division of Applied Psychology, Ontario Institute for Studies in Education, Toronto.

contrast effects inherent in the learning list itself. The present study attempted to eliminate this factor by using learning lists homogeneous for affective tone, but differing from other homogeneous lists in 'pleasantness' or 'unpleasantness'.

Another possible source of confusion, especially in the early studies, is the lack of control over such other factors as meaningfulness, word frequency, association value, word length and so on, which are known to affect learning and retention. Recent studies (e.g. Keppel, 1963; Amster, 1964) have attempted to achieve more control over these variables, and the most important of them were also controlled in the present study.

METHOD

Materials

A pleasant, unpleasant and neutral list of words was constructed in the following way. First, 120 six-letter nouns and adjectives, occurring between fifteen and forty times per million, were taken from the Thorndike & Lorge (1944) general word count. These words were randomly ordered, and mimeographed in the form of booklets. There were two forms of the booklet, differing only in the random order of the words within the booklet. The first page of each booklet instructed subjects to rate the words either for 'pleasantness' or for 'familiarity'. Half of each form of the booklet was rated on each scale.

The booklets were given to 285 students from two different sections of an introductory psychology class. Seventy-six students from one section and sixty-two from the other rated the words on a seven-point pleasantness/unpleasantness (P) scale. Pleasantness was defined in the instructions as 'the degree to which a word arouses pleasant or comfortable feelings in you', and unpleasantness was defined as the converse. Point 1 on the scale was to be given to 'extremely pleasant' words, point 4 to 'neutral' words, and point 7 to 'extremely unpleasant' words. Eighty-two students from the first section and sixty-five from the second rated the words on a seven-point familiarity (F) scale. Familiarity was defined as 'how often you use or encounter these words in your reading or in talking with other people.' Point 1 on this scale was to be assigned to 'unfamiliar' words, point 4 to words of 'average familiarity', and point 7 to 'highly familiar' words. In rating both P and F, subjects were instructed to use the full range of the scales and to rate each word independently of the others. All the ratings from each section were obtained during a regular class period; the ratings from the second section were obtained one day after those of the first section.

Within each of the two sections, the P and F ratings were averaged for each word. Product-moment correlations between the average ratings for the two sections were computed and were found to be $+0{\cdot}989$ ($P < 0{\cdot}001$) for the P scale and $+0{\cdot}911$ ($P < 0{\cdot}001$) for the F scale. Such large correlation coefficients indicate a high degree of reliability in the P and F ratings.

The two P and F mean ratings for each word, one from each section, were averaged, giving a single mean P and F rating for every word. The product-moment correlation between the mean P and F ratings was only $+0{\cdot}087$, indicating virtually no relation between P and F as defined in this experiment and, hence, that the two rating scales must have been measuring two independent and different underlying processes. (The complete list of 120 words, together with their mean P and F ratings, may be obtained upon request from the senior author.)

From the combined list of 120 words, sixteen pleasant (mean P rating $= 2{\cdot}52$; range $2{\cdot}01$–$2{\cdot}96$), sixteen unpleasant (mean P rating $= 6{\cdot}01$; range $5{\cdot}76$–$6{\cdot}45$), and sixteen neutral (mean P rating $= 4{\cdot}02$; range $3{\cdot}51$–$4{\cdot}45$) words were chosen. These three lists are shown in Table 1. Beside each word is its mean P and F rating. As Table 1 indicates, the three lists were equated for rated familiarity, the mean and range of the F ratings being approximately the same for all lists. Equating the lists for rated F was an attempt to control the cluster of interrelated variables including familiarity, meaningfulness, pronounceability and association value (cf. Underwood & Schulz, 1960), which are known to affect free recall learning.

The words in each of the lists in Table 1 were first put into a random order and then seven other word orders were systematically constructed so that no word occurred in a given serial position nor followed any other word more than once in the eight orders. Each subject received a different one of these eight orders on each of eight learning trials. The sequence in which the eight word orders were administered was randomly varied across subjects.

Subjects

Thirty women taking an introductory course in experimental psychology were used as subjects. All were between the ages of 18 and 25, and none had rated the words previously.

Procedure

Each subject learned all three of the lists. The order in which the lists were presented was counterbalanced over subjects, with five subjects learning each of the six possible orders of presentation of the three lists.

Table 1. *Lists of pleasant, unpleasant, and neutral words used in the experiment*

Pleasant			Unpleasant			Neutral		
	Mean *P* rating	Mean *F* rating		Mean *P* rating	Mean *F* rating		Mean *P* rating	Mean *F* rating
Turkey	2·48	4·86	Hatred	6·22	4·47	Sample	3·82	4·91
Circus	2·84	3·80	Horror	5·84	4·33	Runner	3·74	4·35
Luxury	2·16	4·80	Gloomy	5·82	4·42	Cement	4·18	4·31
Melody	2·01	4·82	Fright	5·76	4·84	Helmet	4·20	3·24
Chapel	2·62	4·41	Misery	6·18	4·50	Canvas	3·54	3·59
Marvel	2·72	3·73	Menace	5·86	4·14	Cellar	4·42	4·54
Cherry	2·35	4·28	Famine	6·45	3·42	Kettle	3·51	4·82
Hearty	2·61	3·96	Tragic	6·12	4·83	Hammer	4·22	4·37
Velvet	2·42	4·19	Wicked	5·96	3·87	Trader	4·26	3·15
Sacred	2·81	4·20	Victim	5·76	4·84	Sermon	4·07	4·86
Picnic	2·16	4·24	Coward	6·10	4·17	Jungle	4·45	3·44
Parade	2·96	4·37	Strife	5·83	3·57	Barber	4·07	4·11
Cradle	2·67	3·65	Deadly	6·15	4·37	Ladder	3·86	4·37
Riches	2·43	4·38	Guilty	5·85	4·90	Notion	3·58	4·19
Summit	2·75	3·43	Plague	6·36	3·52	Absurd	4·38	4·74
Heroic	2·27	4·76	Robber	5·85	3·83	Debate	3·99	4·85
Mean	2·52	4·24	Mean	6·01	4·25	Mean	4·02	4·24

Each list was presented for eight trials on a 'Stowe' memory drum. A different order out of the eight word orders described previously was used on each trial. The words were presented at the rate of 2 sec/item. After each trial, subjects were allowed 75 sec in which to write down all the words they remembered. There was a break of 2 min between lists. No information was given to subjects about the nature of the material they would be given to learn. They were told only that there would be eight learning trials on each of three different lists of words.

Results

The mean number of words correctly recalled on each list is shown in Fig. 1 as a function of trials; clearly the pleasant and neutral lists were learned more easily than the unpleasant one, but there was little difference between the neutral and pleasant lists. In order to assess the significance of these results a 'cross-over' analysis of variance (Cochran & Cox, 1957) was made on an arc sine transformation of the number of items correctly recalled. Results of the analysis showed that the differences in recall due to affective connotation of the lists were significant ($F = 27{\cdot}33$; D.F. 2, 56; $P < 0{\cdot}0005$). The order in which the three lists were presented to subjects had no significant effect. There was, however, a significant overall practice effect ($F = 32{\cdot}19$; D.F. 2, 56; $P < 0{\cdot}0005$). Ignoring type of material, the mean number of words recalled

per trial for the first list learned was 11·9, for the second 12·9, and for the third 13·4. This slight, but significant, increase over lists probably reflects a learning-to-learn effect, with general improvement on each successive list. The only other significant effect was due to trials ($F = 184{\cdot}3$; D.F. 7, 560; $P < 0{\cdot}0005$). None of the interactions was significant.

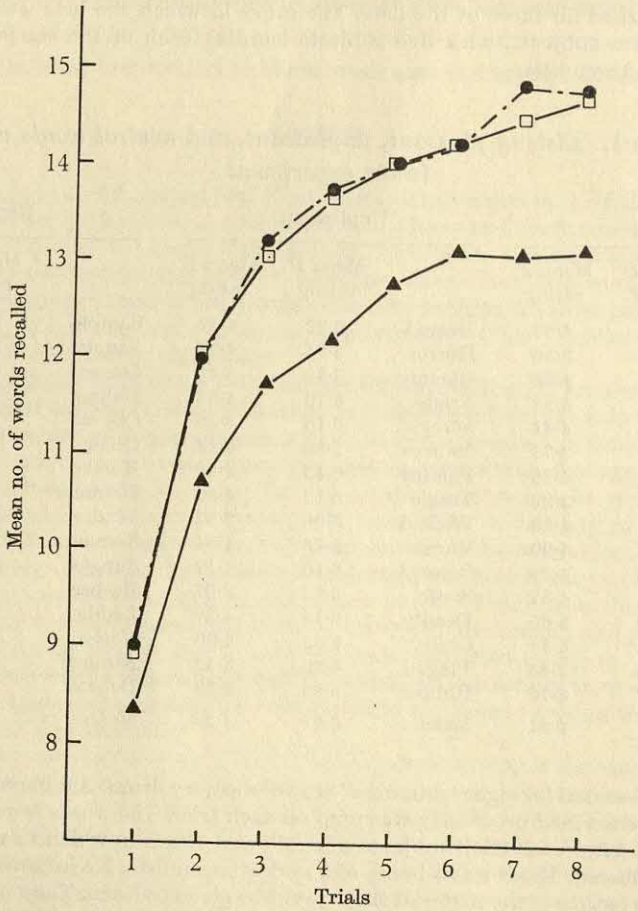


Fig. 1. Mean number of pleasant □—□, unpleasant ▲—▲, and neutral ●-·-● words correctly recalled on eight learning trials.

DISCUSSION

The finding that fewer unpleasant words were learned than pleasant ones is in agreement with many studies investigating the effects of affective factors upon recall. It is the relative position of the neutral list that conflicts with the results of previous studies. The majority of these studies have found recall to decrease in the order: pleasant, unpleasant and neutral. Nevertheless, the differences have often been small and exceptions have not been uncommon. To what, then, can the poorer learning of unpleasant words in the present experiment be attributed? The fact that the words in each list were homogeneous in affective connotation precludes any explanation in terms of contextual or contrast effects. Such factors as word length, frequency,

familiarity (and, therefore, other variables highly correlated with it) were also controlled.

Since there were no differences between the pleasant and neutral lists, it might be tempting to postulate that some sort of defence mechanism was at work, akin perhaps to the phenomenon of perceptual defence. As with perceptual defence, however, a repression hypothesis may not be necessary to explain these results, not at least until other factors can be ruled out. Higher intralist similarity might have accounted for the poorer learning of the unpleasant list. It is true that most studies have shown that intralist similarity of meaning facilitates rather than retards item learning (e.g. Underwood, Runquist & Schulz, 1959). Yet many more intrusion errors, of words similar in meaning or affective connotation to the words in the list, were made on the unpleasant list than on either of the other two lists. The mean number of errors of this type on the pleasant list was 1·66 and on the neutral list 1·13. A t test indicated that this difference was not significant. The mean number on the unpleasant list was 5·50, and was significantly greater than that for the other two lists (both t ratios $> 2{\cdot}10$; D.F. 18; $P < 0{\cdot}05$). There were no errors of confusion, i.e. of words from a previously learned list.

If high intralist similarity was the cause of the poorer learning of the unpleasant list, as the error scores indicate, then the possibility of an interaction between intralist similarity and affective connotation should be investigated further. It has generally been assumed (e.g. Underwood *et al.* 1965) that high conceptual or meaningful intralist similarity facilitates free recall learning. Yet none of the studies investigating intralist similarity has used word lists of negative (i.e. unpleasant) connotation. The assumption may hold, therefore, only for words of neutral or positive (i.e. pleasant) connotation.

Although unlikely, there is the possibility that the present results have been influenced by the effects of still other variables which were not systematically controlled. Two such variables, for example, are 'abstractness' (Gorman, 1961) and 'vividness' (Tulving, McNulty & Ozier, 1965). As with all studies examining the relation between learning and some aspect of the stimulus material, it was impossible to control every variable which might conceivably have affected the learning of the lists. Nevertheless, further refinement of the lists that are used and further control over other variables known to affect learning are going to be necessary before unequivocal results can be presented. This experiment is only an initial attempt to make some sense out of the morass of conflicting results in the literature on the affectivity variable.

This research was supported by Grant no. APT-86 from the National Research Council of Canada. A preliminary version of the paper was presented at the annual meeting of the Canadian Psychological Association in Montreal, June, 1966.

References

Amster, H. (1964). Evaluative judgement and recall in incidental learning. *J. verb. Learn. verb. Behav.* **3**, 466–73.

Balken, E. R. (1933). Affective, volitional, and galvanic factors in learning. *J. exp. Psychol.* **16**, 115–28.

Barrett, D. M. (1938). Memory in relation to hedonic tone. *Arch. Psychol., N.Y.* **31**, no. 223.

Cason, H. (1932). The learning and retention of pleasant and unpleasant activities. *Arch. Psychol., N.Y.* **21**, no. 134.

COCHRAN, W. G. & COX, G. M. (1957). *Experimental Designs*, 2nd ed., pp. 127–33. New York: Wiley.

GORMAN, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *J. exp. Psychol.* **61**, 23–9.

KEPPEL, G. (1963). Word value and verbal learning. *J. verb. Learn. verb. Behav.* **1**, 353–6.

MCGEOCH, J. A. & IRION, A. L. (1952). *The Psychology of Human Learning*. New York: Longmans Green.

MELTZER, H. (1930). The present status of experimental studies on the relationship of feeling to memory. *Psychol. Rev.* **37**, 124–39.

RATLIFF, M. M. (1938). The varying function of affectively toned olfactory, visual and auditory cues in recall. *Am. J. Psychol.* **51**, 695–701.

STAGNER, R. (1933). Factors influencing the memory value of words in a series. *J. exp. Psychol.* **16**, 129–37.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Teachers College, Columbia University.

TULVING, E., MCNULTY, J. A. & OZIER, M. (1965). Vividness of words and learning to learn in free-recall learning. *Can. J. Psychol.* **19**, 242–52.

UNDERWOOD, B. J., EKSTRAND, B. R. & KEPPEL, G. (1965). An analysis of intralist similarity in verbal learning with experiments on conceptual similarity. *J. verb. Learn. verb. Behav.* **4**, 447–62.

UNDERWOOD, B. J., RUNQUIST, W. N. & SCHULZ, R. W. (1959). Response learning in paired-associate lists as a function of intralist similarity. *J. exp. Psychol.* **58**, 70–8.

UNDERWOOD, B. J. & SCHULZ, R. W. (1960). *Meaningfulness and Verbal Learning*. New York: Lippincott.

(*Manuscript received* 5 *October* 1966)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 57–62
Printed in Great Britain

DIGRAM-FREQUENCY AND IMMEDIATE MEMORY

By M. J. A. HOWE*
Department of Psychology, Sheffield University

Digram structure is a determinant of immediate memory for consonant lists. Its effects might operate either before or after recall is required, or both. If digram frequency affects the way sequences are stored, a distracting task may have unequal effects on lists containing digrams of high and low frequency. Two experiments were carried out, which produced evidence suggesting that digram frequency influences the way in which sequences are stored prior to recall being required. When letter sequences were presented between presentation and recall of consonant lists consisting of medium-frequency digrams, recall of the latter was positively related to the digram frequency of the interfering sequences.

It has been found that immediate memory for lists of letters is positively related to the similarity to English of the sequential structure of the letters. Thus Miller, Bruner & Postman (1954) observed that reproduction of eight-letter sequences varied with order of approximation to English, and Baddeley (1964*a*) showed that this result was due primarily to factors effective at a post-perceptual stage in the task. Mayzner & Schoenberg (1964) found that immediate memory for consonant sequences is related to the frequency with which the digrams constituting the sequences occur in English, and Baddeley (1964*b*) used an index based on the relative frequencies of digrams in English to predict successfully the performance of subjects at verbal paired-associate and serial learning tasks.

A step towards understanding how sequential structure in letters influences memory would be to specify the processes within a task at which the variable is directly effective. There are two broad stages at either or both of which digram frequency in letter sequences could affect performance at a memory task. First, digram frequency may affect processes determining the storage of letters: a frequent digram might be stored as a single unit, thus lightening the load on the available capacity for retention by reducing the number of units to be held. Secondly, it is possible that digram frequency may directly affect recall processes, and thus influence retention scores. For instance, Deese (1961) has suggested that when reproduction of a letter sequence structurally similar to English is required, 'construction-at-recall' occurs, whereby a subject uses his knowledge of the language to enhance his chance of correctly guessing those letters in a sequence which could not otherwise be recalled.

The aim of the present study was simply to differentiate between influences effective at recall, and prior to recall. The method used was to introduce a variable which could be reasonably supposed to influence storage processes, but not recall processes. Subjects listened to word lists which differed with respect to digram frequency. Recall of these sequences was sometimes preceded by recall of other items. It was reasoned that the interference with retention caused by such additional recall might well depend on how items were stored, but that effects of construction-at-recall would be relatively independent of such previous interference. Thus it would be useful to determine whether consonant sequences containing digrams of high and

* Now at Tufts University, Medford, Mass., U.S.A.

low frequency differ in the extent to which accuracy of recall is affected by the interfering activity of recalling other items. Such a difference would indicate the likelihood of a disparity in storage between sequences containing digrams of low and high frequency. If recall scores for consonant sequences differing in mean digram frequency were equally affected by interfering activity, these would be no evidence that digram frequency influences the way consonant sequences are stored, and the results would thus be consistent with the hypothesis that sequential structure exerts a direct influence mainly at the time of recall.

Experiment I

Method

Experimental task. Lists containing the nine consonants C D F H L N R S T were presented by tape recorder, one item per second. There were two types of list, (H and L). In the L lists average digram frequency (Underwood & Schulz, 1960) within the trigrams formed by letters 1, 2, 3 ('first trigram') and 7, 8, 9 ('third trigram') was low (mean = 6·25). In the H lists the first and third trigrams contained high-frequency digrams (mean = 207·76). The average digram frequency within the letters presented in list positions 4, 5 and 6 (here called 'buffer trigrams') was of an intermediate value, and was constant between the types of list (mean = 31·15 in L lists, 30·00 in H lists). The nine-consonant sequences were constructed so that the mean digram frequencies for letters in positions 3, 4 and 6, 7 were constant between list types, and any letter appeared equally often in each position.

Subjects and procedure. Eighteen paid undergraduate subjects were tested individually. They listened to twenty H lists and twenty L lists. Recall of ten of each type of list was required in the order of presentation ('forward recall'). After the other lists the required procedure ('reverse recall') was to reproduce first the third trigram, then the buffer trigram, and finally the first trigram. Each combination of list type and recall order occurred five times in each successive twenty lists. Six practice lists preceded the experimental trials, which were presented in four orders, balanced among subjects.

After presentation of a list, one of two lights was illuminated to indicate the required recall order. Subjects dictated their attempts to the experimenter, who scored the number of letters which were recalled in their correct position. Presentation of a list began 5 sec after a subject had completed his attempt to recall the previous list.

Results

The symbols P_1, P_2 and P_3 are used to represent letters 1–3, 4–6, and 7–9 (i.e. first, buffer and third trigrams respectively). R_1, R_2 and R_3 stand for the positions of trigrams in the recall sequence. Thus HP_3R_1, represents an H trigram whose letters occurred in positions 7, 8, 9 in the presentation sequence, and which were the first three letters required for recall. Fig. 1 shows recall as a function of the experimental variables. Since each subject was tested in all conditions. D.F. = 17 for all the succeeding comparisons between means. Among trigrams required to be recalled first, those which were presented in positions 7, 8, 9 (P_3) were more accurately reproduced than those which occurred in positions 1, 2, 3 (P_1) in presentation sequences. This result occurred with both list types ($LP_3R_1 > LP_1R_1$, $t = 12{\cdot}22$, $P < 0{\cdot}001$, and $HP_3R_1 > HP_1R_1$, $t = 7{\cdot}53$, $P < 0{\cdot}001$). In the L lists trigrams recalled last were more accurately reproduced if they were first in the presentation sequence ($LP_1R_3 > LP_3R_3$, $t = 2{\cdot}08$, $P < 0{\cdot}04$), but this was not the case for trigrams occurring in H lists ($HP_3R_3 > HP_1R_3$, $t = 0{\cdot}45$, N.S.).

Comparing recall of trigrams in H and L lists, H trigrams occurring in the first

presentation positions in a list and also recalled first (P_1R_1) were more accurately reproduced than L trigrams presented and recalled in the corresponding positions ($HP_1R_1 > LP_1R_1, t = 3{\cdot}96, P < 0{\cdot}01$). Recall of trigrams presented at the beginning of lists and recalled last did not differ singificantly between list types ($HP_1R_3 > LP_1R_3$, $t = 1{\cdot}19$, N.S.), and there was no difference between list types for trigrams presented last and recalled first. However, H trigrams occurring last in both presentation and recall sequences were more accurately reproduced than L trigrams presented and recalled in the corresponding positions ($HP_3R_3 > LP_3R_3, t = 5{\cdot}88, P < 0{\cdot}001$).

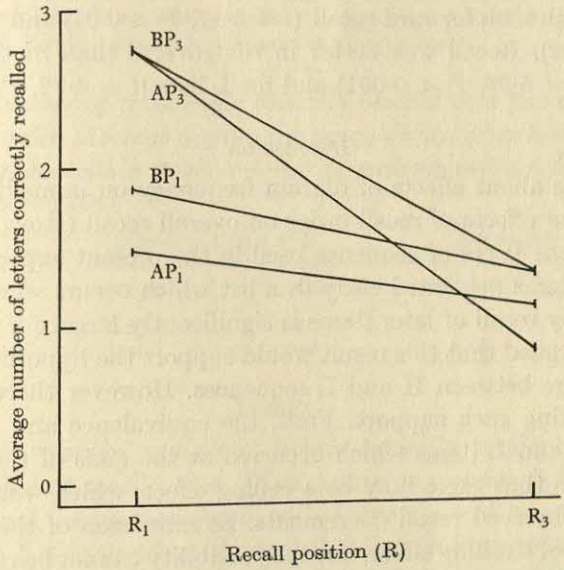


Fig. 1. Average number of letters correctly recalled in nine-consonant lists. P_1 stands for letters 1, 2, 3, and P_3 represents letters 7, 8, 9. L sequences consisted of digrams of low average frequency. In H sequences average digram frequency was high. R_1 and R_3 stand for the first and last three letters to be recalled in a list.

For the reasons explained above a definite storage difference would be indicated if the recall decrement due to interference differed significantly between H and L lists. For trigrams presented at the end of sequences (P_3) there was a greater decrement due to interpolated recall in L lists than in H lists ($LP_3R_1 - LP_3R_3 > HP_3R_1 - HP_3R_3$, $t = 4{\cdot}55, P < 0{\cdot}001$). Thus evidence is provided for a storage difference. However, it is puzzling to note that when a comparison is made between H and L lists for the effect of interpolated recall on reproduction of trigrams occurring at the beginning (P_1) of the nine-consonant sequences, it is found that the comparable decrements again differ between list types, but the direction of the difference is opposite to that observed in the P_3 trigrams ($HP_1R_1 - HP_1R_3 > LP_1R_1 - LP_1R_3, t = 2{\cdot}88, P < 0{\cdot}02$). A smaller $R_1 - R_3$ difference between H and L lists for P_1 than for P_3 items might be attributable to extra 'noise' due to presentation of consonants, but the observation of a significant difference in the opposite direction to that found for P_3 trigrams cannot be simply explained.

Recall of buffer trigrams (P_2), which were always the middle item in the recall sequence, was affected both by the type of list in which they occurred, and by the

order in which the list was recalled. With reverse recall, buffer trigrams occurring in H lists were more accurately recalled (mean = 1·54) than those L lists (mean = 1·28), and, for difference between means, $t = 2{\cdot}72$, $P < 0{\cdot}02$); with forward recall there was a non-significant difference in the same direction (mean = 1·29 in H lists, 1·08 in L lists, $t = 1{\cdot}40$, N.S.). In H lists, recall of buffer trigrams was significantly more accurate with reverse than with forward recall ($t = 2{\cdot}78$, $P < 0{\cdot}02$), and in L lists buffer trigrams were non-significantly recalled when the direction of recall was forward ($t = 1{\cdot}56$, N.S.).

Mean recall of the nine-consonant sequences as wholes was more accurate in H than in L lists, both with forward recall ($t = 3{\cdot}17$, $P < 0{\cdot}01$) and with reverse recall ($t = 2{\cdot}71$, $P < 0{\cdot}02$). Recall was better in the reverse than in the forward order, both for H lists ($t = 5{\cdot}09$, $P < 0{\cdot}001$) and for L lists ($t = 4{\cdot}72$, $P < 0{\cdot}001$).

Discussion

Previous findings about effects of digram frequency on immediate memory have been confirmed, and effects of recall order on overall recall (Howe, 1965) have been found to occur in the types of sequence used in the present experiment. The decrement in recall of items presented early in a list which occurs when reproduction of them is preceded by recall of later items is significantly larger for L lists than for H lists. It has been argued that this result would support the hypothesis that there is a difference in storage between H and L sequences. However there are a number of reasons for qualifying such support. First, the equivalence and near-perfection in recall at R_1, of H and L items which occurred at the ends of their respective lists (HP_3LP_3) indicates that there may be a ceiling effect, which would complicate the interpretation of observed recall decrements. Examination of the data revealed no definitive evidence of a ceiling effect, but the possibility cannot be ruled out, Secondly, as has been mentioned, the decrement which occurred when recall of items presented early in lists was preceded by recall of the items presented later was significantly greater in the H lists than in the L lists, and this result has not yet been explained. A third objection is evident from the recall scores for the buffer trigrams. These scores are related to the digram frequencies of the lists containing them. It is implicit from this that the distraction provided by attending to and recalling other trigrams cannot be regarded as equivalent for H and L lists. Similarly, proactive effects of the presentation of P_1 trigrams on recall of P_3 trigrams may differ between list types.

A second experiment was designed for which the underlying rationale was identical to that of Expt. 1, but which incorporated certain precautions to facilitate unambiguous interpretation of the results. The difficulty level was such that recall never approached 100%, and the items intervening between presentation and recall of a sequence did not differ between list types.

Experiment II

Method

Experimental task. Lists containing the nine consonants used in the previous experiment were presented by ear at the rate of two items per second. There were two types of list. In the first (H lists) the digrams between items 5 and 9 (i.e. 5–6, 6–7, 7–8, 8–9) were of high average frequency (mean = 236), while in the equivalent part of the other (L) lists average digram

frequency was low (mean = 13·4). Average digram frequency within letters 1–4 was equivalent between types of list (mean = 35·3 for H lists, 35·4 for L lists) as was that of the digrams formed by items 4–5 (mean = 9·6 for H lists, 12·6 for L lists). No three letters occurred more than once. Letters were equally distributed between positions in lists.

Subjects and procedure. Fifty-three undergraduate subjects were tested in three groups. The results were analysed together, no systematic inter-group difference being evident.

Each subject heard eight H and eight L lists. Four of each were recalled in a 'forward' order, paralleling that of presentation. With the other ('reverse') order subjects had to recall first letters 5–9, and then letters 1–4. Immediately after presentation of lists, one of two lights showed subjects which recall order was required. Both the first and second eight lists contained two of all the H–L/forward–reverse combinations, arranged randomly. Subjects were allowed 15 sec to attempt written recall on sheets which contained separated spaces for the first four and last five letters in each list. Scoring was as in the previous experiment.

Table 1. *Average number of consonants correctly recalled as a function of type of list (H, L) and of recall order. Average digram frequency within items* 5–9 *is high in H lists, low in L lists. Digrams within items* 1–4 *are of medium average frequency, which is constant between types of list*

	Letters 1–4 (out of 4)		Letter 5–9 (out of 5)	
Recall order	H lists	L lists	H lists	L lists
Forward	2·63	2·34	3·18	2·20
Reverse	2·19	2·09	3·68	3·27

Results

Table 1 shows the average number of letters correctly recalled in Expt. II as a function of list type and recall order. H sequences (i.e. items 5–9 in H lists) were better recalled than L sequences, both with forward recall (i.e. pertinent letters recalled after items 1–4) and when the last five letters were recalled first (reverse recall) ($t = 2·51$, $P < 0·02$ for reverse recall; $t = 7·36$, $P < 0·001$ for forward recall). (D.F. = 52 for this and the succeeding comparisons.) Both H and L sequences were better recalled with reverse than with forward recall ($t = 3·85$, $P < 0·001$ for H sequences; $t = 4·99$, $P < 0·001$ for L sequences). There was a greater recall difference between H and L sequences when they were recalled before rather than after letters 1–4, but the increase was not significant ($t = 1·50$).

With both orders of recall, letters 1–4 were better remembered when they were followed in presentation by H than by L sequences ($t = 3·77$, $P < 0·001$ with forward recall; $t = 2·47$, $P < 0·02$ with reverse recall). Recall order significantly affected retention of letters 1–4 when the sequence 5–9 was H ($t = 2·39$, $P < 0·05$) but not when consonants 5–9 made an L sequence ($t = 0·72$, N.S.). The effects of digram frequency within letters 5–9 on recall of letters 1–4 did not differ significantly between recall orders ($t = 1·21$, N.S.).

The non-significant effect of recall order on the difference in recall scores between H and L sequences may point to a genuine population difference, since $t = 1·50$ approaches the value required for significance at the 0·05 level, by a one-tailed test.

Discussion

The results of Expt. II are consistent with the hypothesis that sequential constraint in word lists influences storage, but on their own offer rather weak support. However, the combined results of Expts. I and II, which are in clear agreement, indicate that digram frequency probably influences recall largely via storage differences, rather than by facilitating correct construction at recall.

It is of interest that recall of consonants 1–4 in Expt. II was influenced by the mean digram frequency of the letters subsequently presented. In other words, the distracting effect of presenting letters 5–9 is (negatively) related to their sequential constraint. This is in accord with a suggestion by Moray (personal communication) that there is a central processing pool of fixed capacity but with interchangeable functions. In Expt. II, using more of the total capacity for perception (and/or initial storage) of items 5–9 leaves less available for storing items 1–4. Since a sequence containing frequent digrams is less redundant than one made up of infrequent digrams, one would predict that coding of H sequences on presentation would require less capacity than coding L sequences, and hence interfere less with retention of items 1–4, leading to the present result.

I wish to thank Prof. H. Kay for useful advice and for providing the facilities for this research, which was done while the author held an S.R.C. Research Studentship.

References

Baddeley, A. D. (1964*a*). Immediate memory and the 'perception' of letter sequences. *Quart. J. exp. Psychol.* **16**, 364–7.

Baddeley, A. D. (1964*b*). Language-habits, S–R compatability and verbal learning. *Am. J. Psychol.* **77**, 463–8.

Deese, J. (1961). From the isolated verbal unit to connected discourse. In *Verbal Learning and Verbal Behaviour*, ed. C. H. Cofer, pp. 11–31. New York: McGraw-Hill.

Howe, M. J. A. (1965). Intra-list differences in short-term memory. *Quart. J. exp. Psychol.* **17**, 338–42.

Mayzner, M. S. & Schoenberg, K. M. (1964). Single-letter and digram frequency effects in immediate serial recall. *J. verb. Learn. verb. Behav.* **3**, 397–400.

Miller, G. A., Bruner, J. S. & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *J. gen. Psychol.* **50**, 129–39.

Underwood, B. J. & Schulz, R. W. (1960). *Meaningfulness and Verbal Learning*. Philadelphia: Lippincott.

(*Manuscript received* 19 *December* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, pp. 63–67
Printed in Great Britain

REPETITION AND IMMEDIATE MEMORY

By E. C. DALRYMPLE-ALFORD
American University of Beirut, Lebanon

Subjects were presented with digit-combinations one digit at a time, in such a manner that each succeeding digit was presented only after all the preceding digits had been repeated by the subject in correct order. The number of digits the subjects could 'carry' in general did not differ from their memory spans. This together with the distribution of serial positions at which repetition broke down is taken as evidence for the view that repetition restores but does not further strengthen memory traces.

Brown (1958) has suggested an explanation of the limit to the span of immediate memory in terms of the decay of memory traces set up during the presentation of the message to be recalled. According to this, when the trace decays beyond a critical level forgetting may occur. Trace-decay may be counteracted by rehearsal during the presentation-recall period, but since simultaneous rehearsal of all the items is impossible, there is a limit to the number of items whose traces may be restored in this way. When the number of items exceeds this limit (i.e. the span of immediate memory) rehearsal will be ineffective in maintaining the complete message, since traces of some of the items will have decayed beyond the critical point before they come to be rehearsed or recalled. According to Brown, rehearsal, in so far as it is mere repetition, restores the memory trace but does not increase its immunity from decay. Any apparent strengthening of the trace during rehearsal is due to (associative) recoding of the message which effectively reduces the memory load. Thus the facilitative effect of repetition on short-term memory reported by Peterson and others (see Peterson, 1966) may not be due to repetition *per se*, but to recoding of the material over successive presentations. If opportunity for recoding is minimized by having, for instance, fast rates of repetition, one would expect no difference in the retention of a set of items that have been repeated many times and that of items that have been presented just once.

Some support for this view comes from the study of McReynolds & Acker (1959) who found no difference in the immediate recall of lists of nonsense syllables presented at very rapid rates either 16 or 32 times. Postman & Adams (1958) found a significant improvement in performance with repetition for intentional but not for incidental learning and this result is also in accord if recoding is considered to be absent in incidental learning because of a lack of set to learn. More recently, Tulving (1966) has shown that mere repetition of material under what may be characterized as incidental learning conditions has no effect on immediately succeeding free-recall learning trials. Like Brown, he suggests that repetition is effective in producing increments in recall only when it permits the recoding or organization of the material into larger subjective units. Further evidence is given here.

Subjects were presented with digit-combinations, one digit at a time, in such a manner that each succeeding digit was presented only after all the preceding digits had been repeated by the subject in the correct order. Thus if the sequence was 72935 . . ., the subject was first presented with 7 and he repeated 7; he was then

presented with 2, and he repeated 7, 2; then with 9, and he repeated 7, 2, 9; and so on until he made an error in his repetition of the items. The maximum number of items on average that the subject could repeat back correctly under these conditions is here called his 'repetition memory span' (RS). With presentation and repetition rates fast enough to prevent recoding, one would expect from the theory outlined above that RS would be equal to conventional memory span (MS) with similar rates of presentation and recall, since, irrespective of the number of repetitions of the items by the subject, the strength of the traces will be the same as with a single presentation. If, on the other hand, repetition does strengthen the memory trace beyond its original level at first presentation, one would expect RS to be greater than MS. Further, with the RS procedure the earlier items in the series should have stronger traces than the later ones since they would have been repeated more often. Hence, when here is a break-down in repeating back the numbers, the point at which the break-down occurs should be later in the series. However, if the traces of successive items are not differentially strengthened by repetition the point of break-down should be unpredictable.

METHOD

Materials. The digit-combinations used were drawn from tables of random numbers with constraints such that neither consecutive repetitions of any digit nor runs of more than two consecutive digits (e.g. 234) were allowed. This was to reduce the likelihood of recoding. The stimuli were recorded on a continuous strip of paper.

Between the successive digits in the RS condition, successions of dots appeared, the number of which corresponded to the number of digits that had been presented to the subject up to that point.

For the MS condition similar paper strips were prepared but with combinations of 4, 5, 6, 7, 8, 9 and 10 digits, followed by a corresponding number of dots to pace immediate recall of the whole sequence.

Procedure. The materials were presented behind a window (see Brown, 1954, for details of apparatus) at a constant rate of about 1 item (digit or dot)/sec in both the RS and MS conditions. In the RS condition after reading each digit aloud as it was presented, the subject had to repeat back all the previously presented digits one at a time in the correct order keeping pace with the dots as they appeared in the window. A trial was terminated when the first error was made.

During the experimental session four estimates of MS were obtained following the procedure described in Woodworth & Schlosberg (1955). Between these MS trials the subject had five trials under the RS condition (15 trials in all), and the experimenter recorded the point at which the first error in repetition occurred and the maximum number of digits that subjects were able to 'carry' without error.

Subjects. Ten university students acted as subjects. They were given preliminary trials until they were familiar with the tasks.

RESULTS

Table 1 sets out the estimates of RS and MS for each subject. The significance levels were obtained from t tests where $t = (\mathrm{RS}-\mathrm{MS})/(\text{S.E. of RS})$. Fig. 1 shows the proportion of occasions on which subjects were able to repeat back correctly sequences of various lengths under both conditions.

The other result of interest is the distribution of the points at which recall (repetition) broke down under the RS condition (see Fig. 2). Data from occasions when less than six or more than nine items were being repeated back are omitted, since there were fewer than eight cases of each (12 % of the data). The Kolmogorov–Smirnov

one sample test' (Siegel, 1956), although it assumes an underlying distribution that is continuous, was used to test departure from equiprobable breakdown at different serial positions. Only in the cases where subjects failed to 'carry' more than five digits ($k = 6$) was the difference significant ($P < 0{\cdot}01$). In all other cases the probability on the null hypothesis was greater than 0·20.

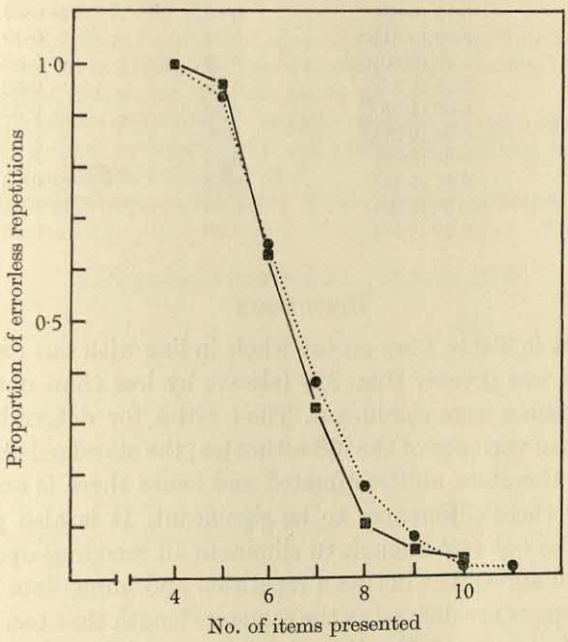


Fig. 1. Proportion of errorless repetitions as a function of number of items presented under repetition memory span (RS) and immediate memory span (MS) conditions: ● · · · ●, RS; ■— ■, MS.

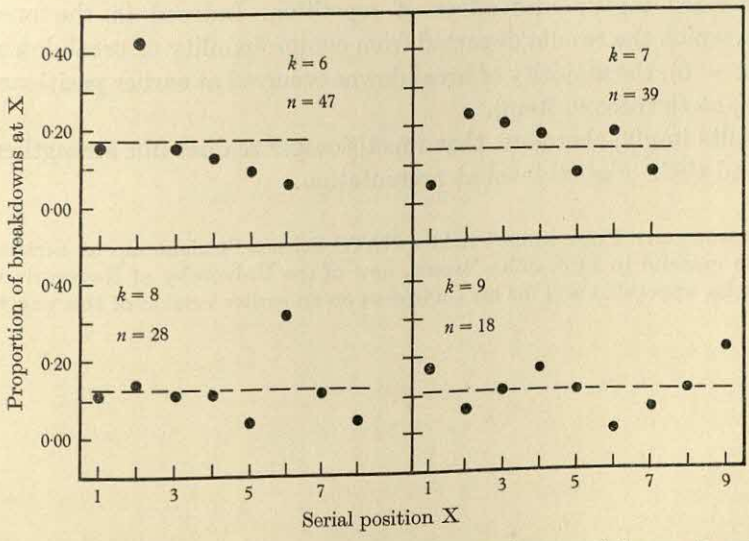


Fig. 2. Proportion of breakdowns in repeating items previously presented occurring at various points in the sequence being called back: k = number of digits to be repeated; n = number of cases on which proportions are based.

Table 1. *Estimated repetition memory span (RS) and immediate memory span (MS)*

Subject	RS Mean	RS S.D.	MS Mean	Significance of (RS−MS) P
1	5·13	(0·83)	5·25	> 0·50
2	5·33	(0·81)	5·50	> 0·40
3	5·73	(0·80)	5·25	< 0·05*
4	5·93	(1·27)	5·75	> 0·50
5	6·00	(1·06)	6·25	> 0·30
6	6·06	(0·96)	5·25	< 0·01**
7	6·27	(1·03)	6·25	> 0·90
8	6·67	(1·35)	6·75	> 0·90
9	7·60	(1·05)	6·75	< 0·01**
10	8·07	(1·83)	8·00	> 0·80

DISCUSSION

The results set out in Table 1 are on the whole in line with our expectation. In six of the ten cases RS was greater than MS (always by less than one digit) and only three of these differences were significant. The t ratios for determining significance took no account of the variance of the MS estimates; the standard errors of the (RS−MS) differences are therefore underestimated and hence there is an increased likelihood of concluding these differences to be significant. It is also possible that the presentation rate was not fast enough to eliminate all recoding operations by these subjects. The overall agreement between repetition and immediate memory spans is even closer if these spans are defined as the sequence length that can be repeated back correctly 50 % of the time (see Fig. 1).

The data illustrated in Fig. 2 are more striking and avoid possible artefacts arising out of arbitrary definitions of memory span. They show little evidence of the generally hypothesized incremental effect of repetition. Indeeed in the one instance examined in which the results departed from equiprobability of breakdown positions (see Fig. 2, $k = 6$), the majority of breakdowns occurred at earlier positions (approximately 40 % at the second item).

These results imply, therefore, that repetition *per se* does not strengthen memory traces beyond their original level at presentation.

This study was carried out while I held a NATO Science Studentship at Birkbeck College, London. I am grateful to Prof. John Brown, now of the University of Newcastle upon Tyne, for the use of his apparatus and for his comments on an earlier version of this paper.

References

Brown, J. (1954). The nature of set to learn and of intramaterial interference in immediate memory. *Quart. J. exp. Psychol.* **6**, 141–8.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.* **10**, 12–21.

McReynolds, P. & Acker, M. (1959). Serial learning under conditions of rapid presentation of stimuli. *Am. J. Psychol.* **72**, 589–92.

Peterson, L. R. (1966). Short-term verbal memory and learning. *Psychol. Rev.* **73**, 193–207.

Postman, L. & Adams, P. A. (1957). Studies in incidental learning. VII: Effects of frequency of exercise and length of list. *J. exp. Psychol.* **56**, 86–94.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill.

Tulving, E. (1966). Subjective organisation and effects of repetition in multi-trial free-recall learning. *J. verb. Learn. verb. Behav.* **5**, 193–7.

Woodworth, R. S. & Schlosberg, H. (1955). *Experimental Psychology.* London: Methuen.

(*Manuscript received* 21 *January* 1967)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 69–75
Printed in Great Britain

INFORMATION PROCESSING RELATED TO STIMULUS NOVELTY AND COMPLEXITY IN A SIGNAL DETECTION PARADIGM

By BERNARD WEINER and PAUL FELDMAN
University of California, Los Angeles

Three experiments are reported which investigated the informational properties of novel and complex stimuli. Irrelevant visual stimuli varying in novelty and complexity were employed as noise in an auditory signal detection task. Signal detection increased over the time of exposure of the visual stimuli, suggesting that the amount of information being processed from a stimulus is a function of its novelty. The judged complexity of the stimulus did not systematically influence signal detection. The investigations employed both within- and between-subjects experimental designs.

In an auditory signal detection experiment, subjects are asked to respond to the presence of a given stimulus. Generally the intensity of the signal and/or amplitude of the noise introduced into the system is varied, and the subjects' responses are, in part, a function of the signal-to-noise ratio. It has proved of heuristic value to view signal detection within the framework of information processing. The organism is considered as having a channel capacity which limits the amount of information that can be processed at a moment in time (Miller, 1956). As the amount of competing information irrelevant to a specified task increases, performance decreases (Broadbent, 1958). Hence one may hold that detection tasks, as well as other performance tasks, are dependent upon the ratio of relevant information (signal) to irrelevant information (noise).

Signal detection has been employed in this study to investigate variables believed to be important in understanding curiosity. Berlyne (1960) maintains that collative variables (e.g., novelty, complexity) are related to the amount of information that is received from a stimulus, and information is held to influence curiosity. To investigate the informational properties of collative variables, visual stimuli varying in novelty and complexity were employed as noise (competing information) in a signal detection paradigm. A relevant auditory signal and the irrelevant visual stimuli were presented simultaneously to the two sense modalities. If the information in a stimulus is related to collative properties, then performance at the auditory detection task should vary as a function of the novelty and complexity of the visual stimuli. The proportion of the subject's information processing capacity utilized during commerce with the complex and novel stimuli was expected to affect efficiency at the detection task.

Berlyne holds that novelty is in part a function of how often and how recently a stimulus has been perceived. Complexity is a function of the physical properties of a figure (e.g. the number of elements, similarity of the elements). Hence, it was hypothesized that detection of auditory signals would increase during exposure of a visual stimulus, and would be negatively related to the complexity of the exposed figure.

EXPERIMENT I

Method

Materials. To approximate a measure of physical complexity, the judged complexity of 107 abstract paintings was determined (see Attneave, 1957; Houston, Garskof & Silber, 1965). Colour slides of the pictures were projected on a screen for 10 sec. After each slide, subjects rated the complexity of the picture against that of a standard picture which was given the value of 100. Ten seconds were allowed to rate stimulus complexity, the standard slide remaining on throughout the experiment. The magnitude estimation task was administered to two groups of male and female students enrolled in an introductory psychology course. For one group ($n = 15$) the order of stimulus presentation was randomized; the stimulus order for the second group ($n = 11$) was reversed. The product-moment correlation between the ratings of the two groups was $r = 0{\cdot}86$. The correlation between the ratings of the males ($n = 12$) and the females ($n = 14$) was $r = 0{\cdot}84$.

Fifteen pictures judged to be most complex ($\bar{X} = 285$), fifteen least complex ($\bar{X} = 64$), and fifteen of medium complexity ($\bar{X} = 144$) were selected for the final experiment. The subjects in the main experiment did not participate in the initial rating task.

Procedure. Subjects were tested individually during the final experimental hour. Initially, an auditory threshold was approximated for each subject by the ascending and descending staircase method (Cornsweet, 1962). The stimulus was a click of 0·1 msec duration produced by a 10 V AC pulse. The stimulus was transmitted into a pair of stereophonic earphones via an attenuator used to adjust the intensity to the threshold value.

The following instructions were then read to the subject: 'On the screen in front of you we are going to flash a series of pictures. They are slides of modern art, and will be shown for about 30 sec. When the picture appears, look at it carefully. Later we will be asking you questions about your preferences among pictures and from your answers we will try to formulate reasons for those preferences. Also, while you are looking at the slides, at times we will present the noise which you heard before. Whenever you hear that sound, please press the button to notify us. Any questions? Remember to concentrate on the picture and the visual images as they appear on the screen.' The reader should note that to maximize information processing from the visual stimuli, attending to those stimuli was established as a primary task.

The forty-five selected stimuli were projected on a screen with the subject seated approximately 5 ft. from the 27 in. × 19 in. image. Two random orders of stimulus presentation were used. Each stimulus was projected for 30 sec with an intertrial interval of 0·75 sec. During the 30 sec the subjects received an equal number of trials with 0, 2, 3, 4, or 6 signal presentations. The signals were unequally interspersed throughout the interval. Signals were presented 135 times per subject, with forty-five signals at each complexity level. The subjects responded by pressing a button which lighted a bulb not visible to them. Responses 2 sec after the signal were considered false.

There was one practice trial with four signals to familiarize the subject with the procedure and to refine further the threshold setting. If the subject responded to more or less than two of the signals, the auditory signal was raised or lowered 1 db. The practice trial was accompanied by a high complexity stimulus.

Subjects. Subjects were fifty-one men and women students.

Results

There was an average of 8·4 false responses per subject. Four subjects accounted for 33% of the false responses. The false responses proved to be unrelated to the complexity or exposure time of the stimulus; they have been neglected in all ensuing analyses.

Fig. 1 illustrates the percentage detection of the auditory signal while viewing pictures of varying complexity over time. The figure shows that for high (H) and medium (M) complexity figures detection increased over the first 24 sec, and decreased during the final 6 sec block of time. Detection of the signal during presentation of a low (L) complexity irrelevant stimulus does not systematically increase. An analysis

of variance revealed a significant main effect attributable to the time of exposure ($F = 22{\cdot}85$; D.F. 4,200; $P < 0{\cdot}001$). The main effect due to stimulus complexity did not approach significance ($F < 1$), nor was the complexity × time interaction significant ($F = 1{\cdot}80$; D.F. 8,400; $0{\cdot}10 > P > 0{\cdot}05$).

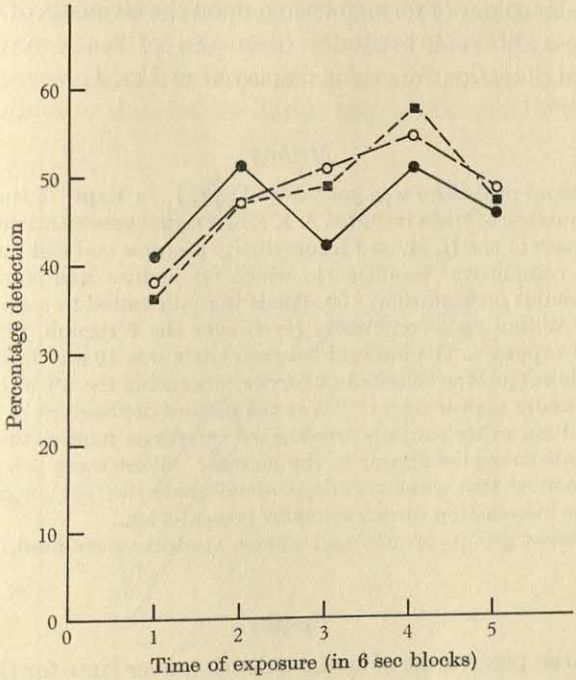


Fig. 1. Signal detection as a function of time of stimulus exposure (novelty) and stimulus complexity: ●——● low complexity, ○——○ medium complexity, ■---■ high complexity.

Discussion

The results indicate that novelty, which decreases as a function of stimulus exposure, affects detection of the auditory signal. More information is therefore processed from novel than familiar figures. This appears to be especially true given figures relatively high in stimulus complexity. Stimulus complexity is not significantly related to detection. During the initial 12 sec of exposure of the irrelevant stimulus there was greater detection during presentation of the low complexity stimulus, as was expected. However, the relationship between detection and complexity was reversed after 12 sec of exposure. At present no interpretation of this reversal can be offered with any reasonable confidence. The drop in detection during the final 6 sec of exposure also is puzzling. The decrement might be caused by the subject's anticipating the onset of the next stimulus rather than attending to the immediate task.

EXPERIMENT II

The thoughts which accompany the perception of an abstract painting may vary along many dimensions. One might respond to the 'feeling-tone' of the picture, primarily concentrating upon its affective qualities. On the other hand, one might analyse the particular colour or form, focusing upon the elements of composition in the artistic work. These different attitudes or modes of reaction conceivably could influence the signal detection functions displayed in Fig. 1.

Method

Procedure. The general procedure was guided by Expt. I. In Expt. II the figures were shown for 40 sec. An equal number of trials included 3, 4, 5, or 8 signal presentations. There were thirty-two trials, eight for each of the H, M, and L complexity pictures and eight additional trials for a control or 'minimal complexity' condition in which no picture was shown. There were two random orders of stimulus presentation; 160 signals were presented to each subject, 40 at each level of complexity. Within each complexity level, over the 8 stimuli, one signal occurred at each of the 40 sec of exposure. The interval between trials was 10 sec. In one condition during that time interval the subjects were asked questions concerning the affect conveyed by the picture, e.g. 'Did the picture appear warm?' 'Was the picture depressing?'

In a second condition, other subjects were asked questions related to the elements in the picture, e.g. 'Did the colour blue appear in the picture?' 'Were there any straight lines in the picture?' It was presumed that these questions would guide the 'set' or manner in which the subject processed the information during stimulus presentation.

Subjects. Two different groups of men and women students were used, twenty-four in each group.

Results

Fig. 2. portrays the percentage of signal detection over time for the four complexity levels when questions related to the affect conveyed by the pictures were asked. An analysis of variance indicated that there was a significant main effect due to time of exposure ($F = 44{\cdot}50$; D.F. 4, 92; $P < 0{\cdot}001$). The main effect of stimulus complexity was not significant ($F = 2{\cdot}18$; D.F. 3, 69; $0{\cdot}10 > P > 0{\cdot}05$), nor was the time × complexity interaction significant ($F = 1{\cdot}65$, D.F. 12, 276; $0{\cdot}10 > P > 0{\cdot}05$).

The pattern of results was very similar for the group which was asked questions related to elements in the pictures. Again there was a highly significant and monotonic increase in signal detection over the time of exposure ($F = 28{\cdot}55$; D.F. 4, 92; $P < 0{\cdot}001$). Neither the level of complexity nor the time × complexity interaction were significantly related to detection ($F < 1$).

Discussion

The results essentially replicate the findings in Expt. I. Again novelty influences detection, while the complexity of the irrelevant stimulus neither significantly nor systematically affects detection. These findings occurred in both the 'element' and 'affect' question conditions. The supposition that the two different sets would differentially influence detection, i.e. occupy differential amounts of channel capacity, was not supported. The data also revealed that the decrement in detection obtained in Expt. I after 24 sec of exposure was no longer exhibited. This lends some credibility to the argument that the decrements displayed were attributable to the subject's

anticipating the next stimulus. In Expt. II there was a 10 sec interval between the offset and onset of the pictures.

A surprising result in this investigation is the increment in the detection of signals as a function of time of 'exposure' in the control condition. One plausible explanation of this result is derived from analysis of the determinants of a response in a signal detection paradigm. In part, the tendency to respond to a signal is a function of the expectancy that the signal will be presented (Deese, 1955). When a picture is shown, the subjects detect more signals over the time of exposure. Hence, they might come to expect more signals later than earlier during any given trial. This could account

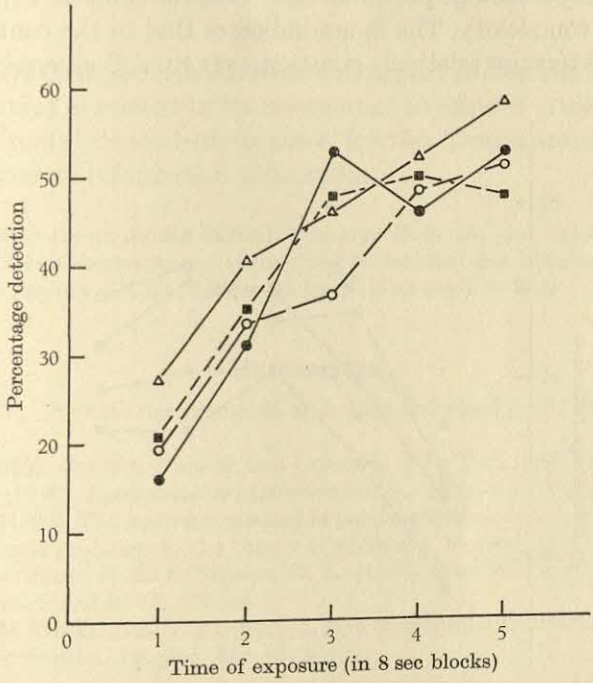


Fig. 2. Signal detection as a function of time of stimulus exposure (novelty) and stimulus complexity, with an 'affective' set △ —— △ control, ● —— ● low complexity, ○— —○ medium complexity, ■- - -■ high complexity.

for the increments in detection over time in the control condition. The above analysis suggests that the responses in a detection paradigm are not independent of one another. Therefore, a within-subjects design might be unsuitable in the present context; responses during the presentation of stimuli at one complexity level could influence responses to stimuli of other complexities. It would be more appropriate to employ a between-subjects design in which different subjects respond to visual stimuli of different complexity; this plan was followed in Expt. III.

Experiment III

Method

Expt. III incorporated aspects of the two previous experiments. Different groups of subjects were shown pictures, all of which were of H, or of M, or of L complexity. A fourth group served as a control group in which no pictures were presented. There were sixteen stimuli in each

condition. Each stimulus was projected for 40 sec; during that time period there were equal numbers of trials with 3, 4, 5, or 8 signals. Over the sixteen trials each subject received 80 signals, two at each of the 40 sec of stimulus exposure. There was a 5 sec intertrial interval. The instructions to the subject were the same as in Expt. I, in which no questions were asked between slides. The instructions were modified in the control condition. The familiarization trial in the H, M, and L complex conditions was accompanied by a picture of H complexity; the pretrial in the control condition was not coupled with a picture.

Subjects. Four groups of fifteen subjects (men and women students) were used.

Results

Fig. 3. shows the percentage detection over time for subjects exposed to different levels of stimulus complexity. The figure indicates that in the control condition the detection of signals remains relatively constant over time. Conversely, when a picture

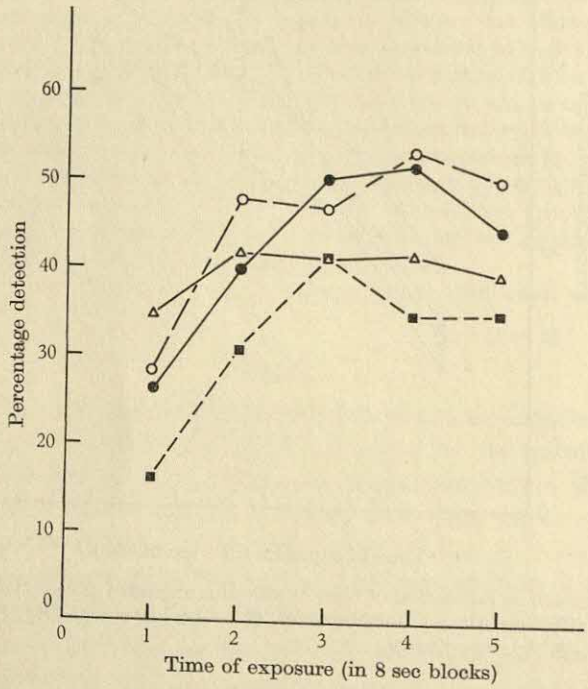


Fig. 3. Signal detection as a function of time of stimulus exposure (novelty) and stimulus complexity with a between-subjects design: △——△ control, ●——● low complexity, ○——○ medium complexity, ■- - -■ high complexity.

was the irrelevant stimulus, detection increased as a function of the time of exposure ($F = 22{\cdot}96$; D.F. 4, 168; $P < 0{\cdot}001$). The figure also reveals that there was least detection of signals in the high complexity condition. However, analysis of variance indicated that the main effect attributable to stimulus complexity does not reach significance ($F = 2{\cdot}86$; D.F. 2, 42; $0{\cdot}10 > P > 0{\cdot}05$). The time × complexity interaction did not approach significance ($F < 1$).

Discussion

Again the data indicate that the amount of noise in this situation decreases with increased exposure of the irrelevant stimulus. That is, more information is processed from a new than from a relatively familiar figure. In contrast to the previous experiment, the results in Expt. III suggest that the complexity of the irrelevant stimulus also affects signal detection. However, detection is not monotonically related to complexity, and the significance of the complexity of the irrelevant stimulus remains to be demonstrated with this paradigm.

Conclusion

In summary, the data provide evidence in support of Berlyne's position that the novelty of a stimulus is related to its informational value. Further, signal detection appears to be a useful dependent measure for the investigation of problems conceptualized in terms of information processing.

The authors wish to thank James Carroll, Penelope Potepan, and Randy Volheim for their aid in conducting these experiments. Computing assistance was obtained from the Health Science Computing Facility, UCLA, sponsored by NIH Grant FR–3.

References

Attneave, F. (1957). Physical determinants of judged complexity of shapes. *J. exp. Psychol.* **53**, 221–7.

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.

Broadbent, D. E. (1958). *Perception and Communication*. London: Pergamon.

Cornsweet, T. N. (1962). The staircase method in psychophysics. *Am. J. Psychol.* **75**, 485–91.

Deese, J. (1955). Some problems in the theory of vigilance. *Psychol. Rev.* **62**, 359–68.

Houston, J. P., Garskof, B. E. & Silber, D. E. (1965). The informational basis of judged complexity. *J. gen. Psychol.* **72**, 277–84.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–97.

(*Manuscript received* 26 *October* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 77–84
Printed in Great Britain

EFFECTS OF AROUSING STIMULATION UPON NOVELTY PREFERENCE IN RATS

By H. CARL HAYWOOD and THEODORE D. WACHS
George Peabody College, Nashville, Tennessee

White rats were habituated to one arm (either black or white) of a Y-maze, then exposed to conditions of intense white noise, electric shock, shock incongruity, or no arousing stimulation, and finally allowed free choice to enter the novel or familiar maze arm. Under all conditions of arousing stimulation, preference for novelty was reduced to a chance level (50 %), while in the non-arousing control conditions, preference for novelty in the two-choice situation was 79 %. The results indicated that strong arousal does cause a decrement in novelty preference, that this decrement occurs regardless of the type of stimulation used in the experiment, and that incongruity may also be sufficient to produce this decrement. A number of theoretical possibilities which might account for these results were explored.

Over the past 15 years a series of experiments has clearly identified a pattern of behaviour, variously termed curiosity (Berlyne, 1951, 1960), perceptual exploratory drive (Montgomery, 1954; Montgomery & Segal, 1955), exploration motivation (Butler, 1953, 1957), novelty seeking (McReynolds & Bryan, 1956; Haywood, 1961, 1962) and manipulation drive (Harlow, Harlow & Meyer, 1950; Harlow & McClearn, 1954). All these various terms refer to the tendency of organisms to expose themselves to new or varied sources of stimulation.

Recently, motivation theorists have begun to specify the environmental conditions which influence preference for novel stimulation (Hebb, 1955; Leuba, 1955; Glanzer, 1958; Hunt, 1963). These theorists have viewed the organism as an information-processing system which seeks an optimal level of stimulation. Below this optimal level, organisms will approach new and varied sources of stimulation; above this optimal level, organisms will seek the familiar in preference to novel stimulation. Research centered around this formulation has shown that the tendency to seek novel stimulation is a relatively strong and persistent motive (Nissen, 1930; Zimbardo & Montgomery, 1957; Mongtomery, 1954), that the degree of novelty seeking within the organism is reliable over time (Butler, 1957), that organisms deprived of stimulation will be motivated to seek out sources of stimulation (Bexton, Heron & Scott, 1954; Jones, Wilkinson & Braden, 1961), that overstimulated organisms will avoid novelty (McReynolds & Bryan, 1956; Haywood, 1962), and that exploratory behaviour is a function of the stimulation value of the environment (Rabedeau & Miles, 1959).

The present study was designed to assess the effect of arousing stimulation upon the preference for novelty. Since Dember (1956) has shown that rats under no particular stress strongly prefer the novel (previously unexplored) endbox in a Y-maze, this situation was used as a control condition. Given an initially high level of novelty preference, one would expect from the foregoing conceptual analyses that greatly increasing the level of arousing stimulation would result in decrements in the preference for novelty, because under these conditions an organism would be beyond the optimal level of arousal.

Two principal questions and one incidental question were asked: (1) Will high levels of arousing stimulation reduce novelty preference? (2) Is novelty preference dependent upon the quality of arousing stimulation? The third question, which can be entertained only incidentally within the present design, arises from an apparent conflict between the conceptual notions of Hebb (1955) and Berlyne (1960), which has been focused by Haywood & Hunt (1963). Hebb (1955) appears to assume that there is an optimal level of *arousal*, above which the organism avoids novelty; however, Berlyne's (1960) assumption is that the optimum is one of *arousal potential*, or *incongruity**, in which case arousal *per se* would not be expected to alter novelty preference. Haywood & Hunt (1963) reported that arousal without the incongruity component (produced by injections of adrenaline) fails to produce the expected decrement in novelty preference. The incidental question asked was: Will a situation designed to produce arousal and incongruity have the same effect on novelty preference as a situation designed to produce arousal without incongruity? Five successive studies were done in order to collect evidence bearing on these questions.

METHOD

The general strategy was to habituate rats to one arm (either black or white) of a Y-maze, then to test their preferences for the familiar maze arm as opposed to the unfamiliar arm in a free-choice trial, with highly arousing stimulation or nothing interposed between habituation trials and the free-choice trial.

Subjects

Subjects were 160 male Wistar rats between the ages of 115 days and 140 days. Each animal was gentled for 2 min daily during the week immediately preceding his participation in the experiment. All the animals were maintained in individual cages in a relatively noise-free colony room at 73°F ± 5°. Food and water were constantly available in the cages.

Apparatus

The basic apparatus was a Y-maze consisting of a startbox and two detachable maze arms. All three components had the same dimensions, being 18 in. long, 10 in. high, and 4 in. wide on the inside dimension. The startbox was painted grey, while one of the endboxes was painted white and the other black. Each of the three boxes had a hinged lid of hardware cloth. The startbox had a metal grid floor which could be independently electrified. The floors of the endboxes were wood, painted the same colour as the sides. The startbox terminated in two guillotine doors, one of clear Plexiglas and the other of aluminium, with the Plexiglas door placed ¼ in. behind the aluminium door. Both types of doors were also placed at the openings of the two endboxes. A grey wooden wedge, an extension of the startbox floor, occupied the floor space between the startbox and the two endboxes. The entire maze was mounted on a grey table which extended beyond the maze in all directions.

Apparatus for white noise stimulation consisted of a white-noise generator, a pre-amplifier, an amplifier, and a 15 in. speaker placed just above the starting end of the startbox and pointed into the startbox at 45°. The intensity of the white noise was adjusted psychologically, rather than physically, using a pilot group of rats comparable to the experimental animals, so that the intensity was just less than that required to produce audiogenic seizures. Intensity of electric shock in the grid floor of the startbox was controlled by a variable resistor, and was also adjusted

* Incongruity seems to be one, but only one, of several kinds of variables included in the general term 'arousal potential', which refers to the various stimulus conditions that usually do (but may not always) produce arousal. (D. E. Berlyne, personal communication.)

psychologically for intensity, i.e., using a pilot group of animals, shock intensity was just sufficient to keep the rats jumping and squealing for 2 min.

Procedure

Each of the five experiments used sixteen experimental animals and sixteen control animals. In all five experiments the training or habituation trials were exactly alike, not only across experiments, but for experimental and control animals alike. On each of three habituation trials the animal was placed in the startbox, and after 15 sec both doors to the startbox and both doors to one or the other of the endboxes were opened simultaneously. The right–left positions of the white and black endboxes were systematically counterbalanced, so that in each group four animals were allowed to enter the white box on the right, four entered the white box on the left, four entered the black box on the right, and four entered the black box on the left. As soon as an animal got all of his body except his tail past a door, the door was gently closed, to prevent retracing of the maze. Each animal was allowed 30 sec in the endbox before being removed, and an additional 1 min in a carrying cage between removal from the endbox and replacement in the startbox for the next trial.

In *Experiment I* trial 4 was the free-choice trial. In the experimental group the animals were placed in the startbox, white noise was turned on at full intensity, and was left on during the 15 sec until all doors were opened, and until the animal had 'escaped' into one of the endboxes. Once an animal had entered the endbox, the white noise was immediately turned off and the doors were closed to prevent retracing. Control animals followed the same procedure without the white noise. In this experiment the right–left positions of the black and white endboxes were the same as in the habituation trials.

In *Experiment II* the procedure followed that of Expt. I exactly, except that after the 15 sec period of white noise in the startbox the startbox doors were opened as well as the aluminium doors of both endboxes, with the endbox Plexiglas doors remaining closed for an additional 15 sec delay at the choice point. This delay was instituted to permit some of the total-disruption effects of the stimulation to dissipate, hence to see whether the animals' choices constituted purely random behaviour (as might be inferred if noise reduced novelty preference in Expt. I but not in Expt. II), or whether the choices reflected a relatively enduring state. At the end of this delay, all doors were opened and the animal was free to make his choice of endboxes. Control animals followed the same procedure, except that the white noise was not present.

Experiment III followed the same procedure as Expt. II (including the 15 sec delay at the choice point), except that the right–left positions of the black and white endboxes were reversed between habituation trials and the fourth or free-choice trial. In this experiment an animal who had found the white endbox on the right during the habituation trials would now find the white endbox on the left at the free-choice trial. This procedure was instituted to check on the confounding of black–white preference with position preference.

In *Experiment IV*, after three habituation trials, the endboxes were removed and experimental animals were given 15 sec of shock in the startbox on each of three shock trials before the endbox doors were opened and the animal was allowed to run out on to the grey table. Control animals in this group were simply placed in the unelectrified startbox to remain for 15 sec and then released to run out on to the grey table. Following these three trials the endboxes were replaced in reverse position, so that an endbox that had been on the right during habituation trials would be on the left on test. The test trial (trial 4) was the same both for experimental animals and for control animals. On the test trial, each animal was placed in the startbox without electric shock, where he remained for 15 sec before the endbox doors were opened. The absence of electric shock where shock had previously been experienced in the startbox constituted the 'incongruity' manipulation. For an additional 15 sec the animal remained at the choice point with both aluminium endbox doors open but with the Plexiglas endbox doors closed. At the end of this 15 sec delay period, both endbox doors were opened and the animal was free to make his choice.

The procedure for *Experiment V* was exactly the same as that for Expt. IV, except that on the test trial shock was again present for the experimental animals. In combination with Expt. IV, this permitted a comparison of the immediate effects of shock with the effects of expectation of shock.

The dependent data were the number of animals in each group who entered the novel or familiar endbox on the test trial. Since both endbox doors were open on the test trial, an entry was scored only when an animal's whole body except the tail had passed the guillotine doors.

RESULTS

In Expt. I, in which intense white noise was administered on the free-choice trial for experimental animals but not for control animals, eight of the sixteen animals in the white-noise conditions entered the novel endbox, and eight entered the familiar endbox. Of the sixteen animals in the control condition, fourteen entered the novel endbox while only two entered the familiar endbox. An exact test on the resulting fourfold table gives $P < 0{\cdot}027$ (one-tailed test).

The only change in Expt. II was the introduction of a 15 sec delay at the choice-point. In this experiment nine of the sixteen animals in the white-noise condition entered the novel endbox, while seven entered the familiar endbox. In the control condition fifteen animals entered the novel endbox while only one entered the familiar endbox. An exact test on this fourfold table gives $P < 0{\cdot}019$ (one-tailed test).

As the cell frequencies were highly similar in both Expts. I and II, it was apparent that the delay procedure had little effect on the animal's behaviour. For this reason the results of Expts. I and II were collapsed into a single fourfold table and analysed by means of a χ^2 test. The resulting frequencies are shown in Table 1, for which $\chi^2 = 11{\cdot}13$ (D.F. 1; $P < 0{\cdot}001$). Hence the arousing stimulation provided by intense white noise did result in a significant reduction in the preference for novel stimulation.

Table 1. *Frequencies of novel and familiar endbox entries under white noise and control conditions (Expts. I and II)*

	Novel choice	Familiar choice	n
White noise	17	15	32
Control	29	3	32

In Expt. III the position of the endboxes was switched between the last habituation trial and the test trial. Under these conditions five animals in the white-noise condition entered the novel endbox while eleven entered the familiar endbox. In the control condition, nine animals entered the novel endbox while seven entered the familiar endbox. It is apparent, from these frequencies, that the white noise again reduced the novelty preference of the experimental animals. For the control animals, however, the switching of endbox positions seemed to result in a competition between stimulus novelty preference and position novelty preference, a distinction which has been made in previous research on the relative strength of these competing preferences (Walker, Dember, Earl & Karoly, 1955). This competition reduces the preference of the control animals for the stimulus novelty choice and increases the preference of the experimental animals for the position novelty choice in a manner leading to choice frequencies differing from the first two experiments. As a function of this conflict the frequency table shows non-significant differences ($P > 0{\cdot}05$).

Both Expts. IV and V involved the use of electric shock as arousing stimulation rather than white noise. In Expt. IV, in which shock in the startbox was interposed between habituation and test, but in which no shock was given in the startbox on

the test trial, eight of the sixteen shock animals entered the novel endbox on test, while eight entered the familiar endbox. In the control conditions, in which no shock was interposed, twelve animals went to the novel endbox and four to the familiar. An exact test on the resulting fourfold table gives $P > 0{\cdot}10$.

In Expt. V shock in the startbox was given both on the intervening trials and on the test trial itself. In the shock condition the novel–familiar split was again eight and eight, while in the control condition thirteen animals chose the novel endbox and three chose the familiar. An exact test on this fourfold table gives $0{\cdot}10 > P > 0{\cdot}05$.

Table 2. *Frequencies of novel and familiar endbox entries under electric shock and control conditions (Expts. IV and V)*

	Novel choice	Familiar choice	n
Shock	16	16	32
Control	25	7	32

Again, the similarity of the cell frequencies in Expts. IV and V indicates that the presence or absence of shock on test trials had little differential effect on choice behaviour. Therefore the frequencies from the two experiments were again collapsed into a single table. This procedure had the additional advantage of overcoming the effects of the position preference–stimulus preference competition in these two experiments by yielding a larger sample for analysis. The frequencies for these two experiments are shown in Table 2. A χ^2 analysis on this fourfold table gave $\chi^2 = 5{\cdot}49$ (D.F. 1; $P < 0{\cdot}02$), indicating that electric shock also serves to reduce the preference for novel stimulation in animals exposed to this shock, as contrasted with animals not exposed to shock.

Discussion

The first question asked of these data was: Will high levels of arousing stimulation reduce novelty preference? The answer to this question is unequivocally affirmative. Both electric shock and intense white noise produced decrements in the preference of rats for novel visual stimulation. It must be noted, however, that the reductions in novelty-seeking behaviour attributable to intense arousing stimulation can be interpreted only as reductions in novelty preference, not as novelty avoidance, since in the five experiments reported here the proportion of choices of the novel endbox never falls significantly below 0·50 in a two-choice situation. Montgomery & Monkman (1955) report similar results using electric shock in a complex maze. These authors interpreted their results as evidence that novel stimulation arouses both an exploratory drive and a fear drive, leading to an approach–avoidance conflict. The authors postulated that intense stimulation, e.g. shock, would raise the avoidance or fear gradient, leading to novelty avoidance. The fact that the present data show a reduction in novelty preference, rather than an induction of novelty avoidance (for which one would wish to see significantly less than 50% choice of the novel stimulus in a two-choice situation), fails to lend support to the interpretation of Montgomery and Monkman. The results do, however, lend some support to the conceptual analysis

presented earlier, i.e. the view that there is an optimal level of arousal, below which the organism will seek novelty, and above which the preference for novelty will be reduced. An alternative hypothesis is that the stimulated animals were simply not making a choice at all. Such a situation might result from the intense stimulation serving to 'erase' the memory traces of the habituation trials, or from total disruption of organized behaviour.

The second question asked was whether the reduction in novelty preference was dependent on the quality of arousing stimulation. The observed reduction in novelty preference in the experimental groups was about the same regardless of whether the animals received shock, white noise or shock incongruity. In all cases of high stimulation there was a sharp decrement in novelty preference. These differences would probably have been even clearer had the endboxes not been switched in the shock conditions and in Expt. III, thus setting up a competition between stimulus preference and position preference which reduced the apparent preference of control groups for novel endboxes. That such a competition exists has been shown by Walker and his associates (Walker *et al.*, 1955) who concluded that while stimulus novelty is most important in governing choice behaviour a definite proportion of choices was shown to be due to position novelty instead. In previous studies (Dember, 1956; Thompson & Higgens, 1958) stimulus and position preference were confounded with each other, and from these results and the behaviour of the control animals in Expts. I and II, where these preferences were also confounded, it is apparent that the control animals in Expts. III, IV and V would have shown much more traditional 'novelty-seeking' behaviour had not this stimulus–position competition been established in these experiments. These results, especially those of experiment three, have led to the suggestion that the decrement in novelty preference under conditions of intense stimulation and stimulus–position preference may well be due to the intense stimulation causing the position preference to become predominant in the preference hierarchy over the normally more important stimulus preference (P. L. Newbigging, personal communication). That this hypothesis has not been raised before may in part be due to the fact that in all other previous studies using intense stimulation, stimulus and position preference have been confounded rather than placed in competition as in the present study.

The third, and incidental question was: Will a situation designed to produce arousal and incongruity have the same effect on novelty preference as a situation designed to produce arousal without incongruity? The evidence on this question is certainly quite meagre. The incongruity manipulation is somewhat confounded by the possibility of residual shock effects, in spite of the fact that no shock is present on the choice trial. However, it is notable that the absence of electric shock in a situation in which shock had been associated with the available cues (Expt. IV) produces exactly the same decrement in novelty preference as is observed in the condition in which shock is again present on the choice trial (Expt. V). These parallel results indicate that while incongruity may not be a necessary condition for altering the direction of behaviour with respect to novel stimulation, it may well be a sufficient condition; however, because of the possibility of physiological residual shock effects in our procedure, it is apparent that the conclusive experiment to answer this question has yet to be done.

The data from the control groups taken alone provide a faithful replication of Dember's (1956) results. Dember found that seventeen of his twenty rats (85%) chose to enter the novel endbox on his free-choice trial. In the present study sixty-three of eighty animals in the unstimulated condition (comparable to the Dember situation), or 79%, chose the novel endbox. Excluding the one deviant situation (Expt. III, in which the control split was 9–7), 84% of the sixty-four remaining control animals chose the novel endbox, an even more striking resemblance to Dember's proportion of 85%.

Further research on the behaviour-directing effects of both arousal and incongruity is desirable. The present study has, however, provided evidence that: (1) strong arousing stimulation does result in a decrement in the preference for novel stimulation; (2) the effects of intense auditory stimulation do not appear to be different from the effects of electric shock, a conclusion which raises the possibility that any arousing stimulus which is sufficiently intense is capable of producing decrements in novelty preference; (3) there is a suggestion that while incongruity is not necessary it may be sufficient to alter the direction of behaviour with respect to novel stimulation. That the above stimulus conditions produce decrements in the preference for novelty is apparent. The underlying reasons for this decrement are still unclear, though there are a number of hypotheses which might be raised to account for this phenomenon, of which five possible explanations were mentioned in this paper.

Parts of this research were supported by a research grant (2M–2529) from the National Institute of Mental Health, U.S. Public Health Service. Both authors received the assistance of two additional grants (HD–43 and HD–973) from the National Institute of Child Health and Human Development. The authors are grateful to Dan Davis for help in subject running and apparatus construction.

References

Berlyne, D. E. (1951). Attention to change. *Br. J. Psychol.* **42**, 269–78.

Berlyne, D. E. (1960). *Conflict Arousal and Curiosity*. New York: McGraw-Hill.

Bexton, W. H., Heron, W. & Scott, T. H. (1954). Effects of decreased variation in the sensory environment. *Can. J. Psychol.* **8**, 70–6.

Butler, R. A. (1953). Discrimination learning by rhesus monkeys through visual exploration motivation. *J. comp. physiol. Psychol.* **46**, 95–8.

Butler, R. A. (1957). The effect of deprivation of visual incentives on visual exploration motivation in monkeys. *J. comp. physiol. Psychol.* **50**, 177–9.

Dember, W. N. (1956). Response by the rat to environmental change. *J. comp. physiol. Psychol.* **49**, 93–5.

Glanzer, N. (1958). Curiosity, exploratory drive, and stimulus satiation. *Psychol. Bull.* **55**, 302–15.

Harlow, H. F., Harlow, Margaret K. & Meyer, D. R. (1950). Learning motivated by a manipulation drive. *J. exp. Psychol.* **40**, 228–34.

Harlow, H. F. & McClearn, G. (1954). Object discrimination learning by monkeys on the basis of manipulation motives. *J. comp. physiol. Psychol.* **47**, 73–6.

Haywood, H. C. (1961). Relationships among anxiety, seeking of stimuli, and level of unassimilated percepts. *J. Personality* **29**, 105–14.

Haywood, H. C. (1962). Novelty-seeking behavior as a function of manifest anxiety and physiological arousal. *J. Personality* **30**, 63–74.

Haywood, H. C. & Hunt, J. McV. (1963). Effects of epinephrine upon novelty preference and arousal. *J. abnorm. soc. Psychol.* **67**, 206–13.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychol. Rev.* **62**, 243–54.

HUNT, J. McV. (1963). Motivation inherent in information processing and action. In *Cognitive Factors in Motivation and Social Organization*, ed. O. J. Harvey. New York: Ronald Press.

JONES, A., WILKINSON, H. J. & BRADEN, I. (1961). Information deprivation as a motivational variable. *J. exp. Psychol.* **62**, 126–37.

LEUBA, C. (1955). Toward some integration of learning theories: the concept of optimal stimulation. *Psychol. Rep.* **1**, 27–33.

MCREYNOLDS, P. & BRYAN, J. (1956). The tendency to obtain new percepts as a function of the level of unassimilated percepts. *Percept. mot. Skills* **6**, 183–6.

MONTGOMERY, K. C. (1954). The role of the exploratory drive in learning. *J. comp. physiol. Psychol.* **47**, 60–4.

MONTGOMERY, K. C. & MONKMAN, J. (1955). The relation between fear and exploratory behaviour. *J. comp. physiol. Psychol.* **48**, 132–6.

MONTGOMERY, K. C. & SEGAL, N. (1955). Discrimination learning based upon the exploratory drive. *J. comp. physiol. Psychol.* **48**, 225–8.

NISSEN, H. (1930). A study of exploratory behavior in the white rat by means of the obstruction method. *J. genet. Psychol.* **37**, 361–76.

RABEDEAU, R. & MILES, R. (1959). Response decrement in visual exploratory behaviour. *J. comp. physiol. Psychol.* **52**, 364–7.

THOMPSON, W. & HIGGENS, W. (1958). Emotions and organized behavior: experimental data bearing on the Leeper–Young controversy. *Can. J. Psychol.* **12**, 61–7.

WALKER, E. L., DEMBER, W., EARL, E. W. & KAROLY, A. J. (1955). Choice alternation. I. stimulus *vs* place *vs* response. *J. comp. physiol. Psychol.* **48**, 19–23.

ZIMBARDO, P. & MONTGOMERY, K. (1957). The relative strengths of consumatory responses in hunger, thirst, and exploratory drive. *J. comp. physiol. Psychol.* **50**, 504–8.

(*Manuscript received* 26 *May* 1966)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 85–92
Printed in Great Britain

DEUTSCH'S MODEL APPLIED TO A CONDITIONAL MAZE

By D. H. HOLDING* and D. PRITCHATT
University of Leeds

Rodents (mice) were rewarded for running over, but not for running under, an obstacle giving access to a final common path. On the basis of Deutsch's postulates it was predicted that the proportion of correct responses during test trials should depend upon the distribution of reinforcement during training trials; the proportion being high following block reinforcement, low following alternated reinforcement and intermediate in a procedure used by Sheldon. These predictions were upheld throughout testing in a maze with a long final path, but only held good for the early test trials in a shorter maze for which predictions are equivocal.

Sheldon (1964) showed that rats 'can choose, on the basis of reinforcement, between different ways of reaching the same goal'. In an apparatus affording pathways over or under an obstacle, leading to a common goal box, fourteen out of fifteen animals reached a stringent criterion in learning to choose whichever pathway was consistently reinforced. These results appear to pose a problem for Deutsch's (1960) theory of behaviour, which is interpreted by Sheldon as predicting that an animal will be unable to choose between two experienced pathways of equal length.

The basis of this interpretation is presumably a simplified version of Deutsch's system, which is described as essentially tropistic. The system represents the neural structure underlying behaviour as a network of links, to which are connected perceptual analysers and elements of the motor system. Learning a maze, consists of lowering the resistance to the flow of excitation down a sequence of links corresponding to the correct pathway; motivational excitation flowing down such a row of links will be diverted to the motor system in the correct order by the stimulation of appropriate analysers at the choice points. The system is designed to 'home' on to those cues represented by analysers which at any stage receive the greatest excitation. Therefore, if the pathway from a common source of excitation first bifurcates and then re-unites at a choice point, it may be argued that the two branches will share the available excitation. The neural maze corresponding to this situation is shown as Fig. 1 (*a*). Hence it is concluded that an animal at choice-point X will be activated equally by the cues of paths O and U and will be unable to make a consistent choice between the two alternatives.

Sheldon's (1964) conclusion may be challenged in two ways. In the first place, it is not entirely clear that the short goal box of the maze used did afford a final common sequence of analysers. The floor cues confronting an animal after surmounting the central obstacle are not identical with the predominantly ceiling cues available to an animal emerging from underneath the obstacle. The two paths only coincide at the food cup, which need not be visually detected. A more definitive test would seem to be provided by arranging a longer final path containing distinctive cues. In the present experiment, therefore, behaviour in a smaller replica of the Sheldon maze was compared with behaviour in a version which included an extension of the final path at right angles to the original.

* Now at University of Exeter.

Even with an unambiguous final path, however, it seems that Deutsch (1960) is able to predict differential learning of the two paths. There is no reason to believe that the excitation traversing paths O and U must be equal, since the resistance of the two paths will come to differ as a result of the training procedures in a manner determined by the schedule of reinforcement. This prediction results from consideration of the effect upon Deutsch's system of postulate 12, which states: 'Each link has a certain "threshold" of excitation. Because of this each link will pass on less excitation than it receives. This "threshold" of excitation of a link will be lowered each time a link is switched off. The lowering of the "threshold" will be *proportionate to the intensity of excitation present in the link at the time.*'

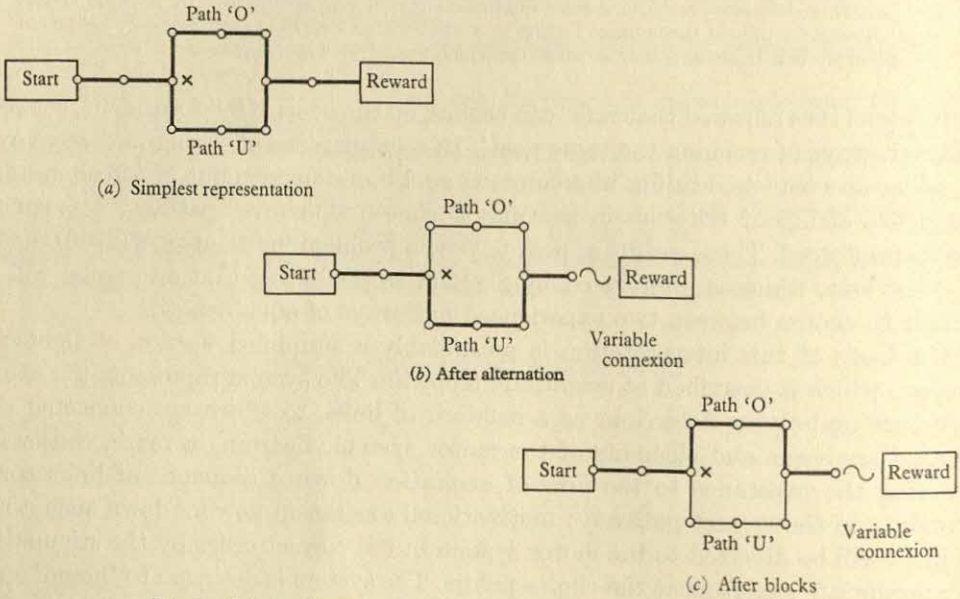


Fig. 1. Schematic neural pathways in terms of Deutsch's system.

The case is first considered in which the animal is forced to alternate between path O, with contingent reinforcement, and path U which leads to no reward. The first run via path U establishes a connected row of links by the normal mechanism of contiguity (postulate 8), stopping short of the primary link which initiates drive excitation. The next run, via path O, establishes a parallel row of links converging with the first row and, since the animal is reinforced, culminates at the primary link. As in Fig. 1 (*a*), there are now two pathways of equal resistance. However, during the next alternation via path U excitation from the primary link is present during stimulation of the analysers and (postulate 12) path U is strengthened. Since no reinforcement is given, the final section of the neural pathway increases in resistance (postulate 11) and less excitation will flow. Thus the connexions through path O will be little affected by the next alternation, although the consequent reinforcement will again reduce the resistance from the primary link. This will benefit path U on the next run. Iteration of the same sequence of events will gradually strengthen path U with respect to path O until the situation shown at Fig. 1 (*b*) obtains.

The unexpected prediction is thus derived that, with *alternation*, reinforcing path O

results in a tendency to choose path U. However, the same prediction is not appropriate when experience of paths O and U occurs in *blocks* of trials. In the case where blocks of O runs precede blocks of U runs prediction is ambiguous, since the quantitative effects of postulates 11 and 12 are unknown. The opposite case, where blocks of U runs precede blocks of O runs, entails the clear superiority of path O. Each of the blocks of U runs will make some connexion by contiguity, but the contributory effect of excitation will be absent. On the first of the blocks of O runs the connexion to the primary link will be made, excitation will begin to flow and, as shown in Fig. 1(*c*), successive runs will establish a relatively strong pathway via O.

The procedure that was used by Sheldon appears to be intermediate between these two cases. Animals alternatively chose and were forced, each forced response being in the direction opposite to the previous free choice. In the early stages of learning this tends to lead to double alternation: an animal forced to O chooses O, is forced to U and next chooses U, is forced to O and so on. The results of this procedure should then fall between those expected under alternation and those expected under block training; in fact, with the technique of opposite-forcing the level of alternation cannot fall below 50 %. The fact that some differentiation between paths took place, but that more than sixty trials were required to achieve it, appears to be consistent with the above analysis.

The three schedules were compared in the present experiment in the expectation that alternation would produce least or opposite learning, that block trials would favour responses in the reinforced direction and that results from the Sheldon procedure should lie between these extremes.

Method

Subjects

The animals were mice (*Mus musculus*) of both sexes, from black, chocolate, cinnamon, blue dilution and albino varieties. They were between 3 and 4 months old at the time of the experiment and naïve but for normal handling. Litters were housed in single-sex cages, from which animals were allotted in approximately equal proportions to the experimental conditions.

Apparatus

A wooden maze with Perspex sides was constructed to half the scale of that used for rats by Sheldon (1964), together with a removable extension; both are depicted in Fig. 2. The start box, and extension arms were 2·5 in. square in section and were all 5 in. long. The centre compartment, which contained the 1·75 in. square block, was 5 in. long and deep and was accessible for cleaning via a removable section. Sliding food cups were positioned underneath a $\frac{7}{16}$ in. diameter hole in the floors of the goal box and extension arms, the centres being located 1·25 in. from each end door. Hardboard screens of 5 in. × 11 in. were positioned as shown to obviate visual cues from the food cups. Sliding Perspex doors were located at the start and at the end of each extension arm and sliding metal shutters at the top and bottom of the centre block, the entrance to the goal box and the entrance to each extension arm.

In addition, a pre-experimental straightway was used. This consisted of 5 in. × 3·5 in. start and goal boxes, linked by a 3 ft. 6 in. runway, all 3 in. high.

Design

The main comparisons were between the alternation, block and Sheldon schedules, and between the Sheldon pattern (short) maze and the (long) version incorporating the extended final path. A total of forty-eight animals was used, in groups of eight on each of the six resulting

conditions. Of the eight animals in any group, four were male and four female. On the short maze it was possible to emerge to the right or the left of the goal box, while on the long maze it was possible to employ either the right arm or left arm of the extension piece; accordingly, half of the animals of each sex in each experimental condition were run to the right, and half to the left.

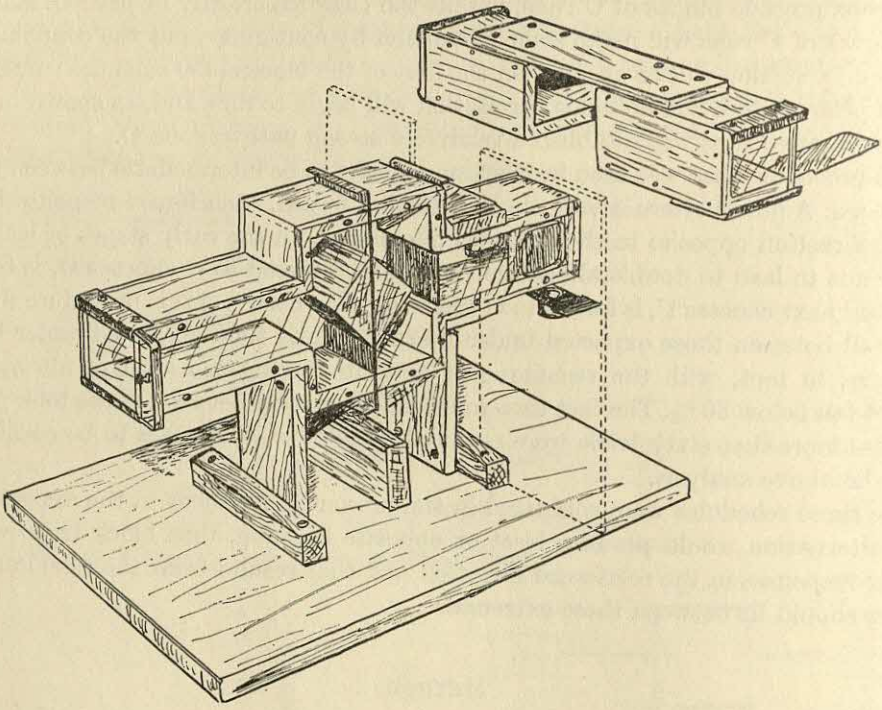


Fig. 2. The conditional maze apparatus.

Procedure

Two days before the experimental period each animal was starved for approximately 12 hr and was then given ten runs in the straightway, in each case reinforced for 30 sec in the goal box. This procedure familiarized the animals with food deprivation, with experimental handling and with timed feeding outside the home cage. Further, the recorded runway times and observed behaviour could be used to eliminate from the experiment those animals who were excessively reluctant or grossly deviant. Nine animals were eliminated in this way; no animals were eliminated on the basis of their behaviour in the subsequent main experiment, except during the first test trial for the reasons noted below.

Before the main experiment animals were starved for an average of 15 hr. On removal from the home cage, each animal was placed in a side box for approximately 3 min. and was then put through the experimental procedure.

On the basis of pilot trials with a dozen animals, the testing period for each animal was divided as follows: (1) a single free test run to ascertain initial preference for passage over or under the block; (2) twenty forced-choice training runs on the appropriate schedule; (3) twenty free runs to test the results of training, the ten final runs being primarily intended to monitor any changes due to learning under test.

Each animal thus completed forty-one runs, of which twenty were forced; all runs were completed in a single session. At the end of each run, and whether or not reinforcement was administered, each animal spent 30 sec in the goal box; he was then transferred to a side box (not the home cage) for 1 min rest before the next run.

Reinforcement was always contingent upon passage *over* the central block. The original intention had been to reward animals for choices opposite to their first preference, but the pilot

trials revealed a strong tendency to pass under the block and to preserve homogeneity it was decided to eliminate any animals initially choosing 'over'. Only three of fifty-one animals were thus eliminated. In each case the choice of 'over' was made hesitantly and after vacillation, in contrast to the apparently firmer choice behaviour of most 'under' animals.

After the test for initial performance, all animals were run with choices forced by closing the shutter on the opposite side of the central block, according to the appropriate schedule:

Block: a series of ten runs 'under' followed by ten runs 'over'; then followed the two sets of ten free test runs.

Alternation: a sequence of twenty over/under alternations, again followed by the free choice runs.

Sheldon: a forced run 'over' was followed by a free choice, the pathway opposite to this next being forced, thus initiating an alternating sequence of forced and free runs which continued to the forty-first and final trial.

During each run all the appropriate doors were closed behind the animal. On reinforced runs a pellet of the animals' normal food (Oxoid diet 41 B) was made available in the food cup. Any crumbs and debris were brushed out between trials. The sliders were greased, but to obviate any information from residual auditory cues the sliders were left sometimes in, sometimes out and sometimes halfway, until the experimenter had observed the direction of choice. Normally four animals were run on each experimental day. The order in which animals in various conditions were run was recorded, as were the deprivation times. These factors were approximately counterbalanced against conditions, although it was impossible to adhere to any rigid scheme of testing.

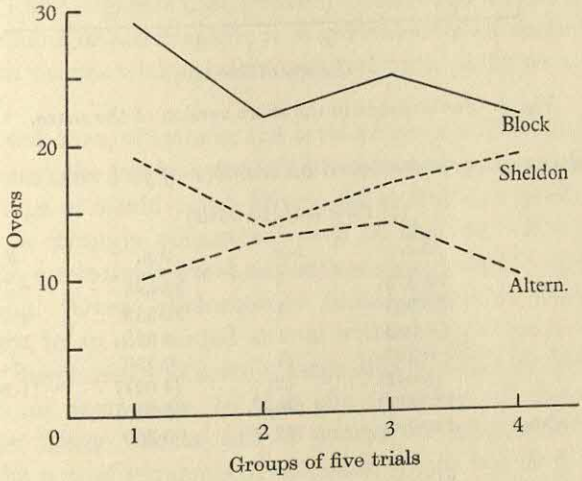


Fig. 3. Performance during the four quarters of the testing period of animals in the long version of the maze.

Results

Numbers of responses in the 'over' direction have been plotted against successive quarters of the free-choice test runs in Figs. 3 and 4. In the long maze the predicted order of 'overs'—block, Sheldon, alternation—was maintained throughout the test trials. Further, since eight animals in the course of five trials could produce a maximum of forty 'overs', the indifference point lies at twenty. On this basis, the block schedule animals predominantly passed over, while the alternation animals passed under, the central obstacle. Behaviour in the short maze appears to vary with the stage of training. The predicted order obtained at the outset, but a depression in the block group scores and a rise in the Sheldon group scores supervened.

Results for the two mazes were pooled for analysis of variance, the earlier and later groups of ten test trials being analysed separately. Both analyses are shown in Table 1. During the first half of the test trials, the differences between both the schedules of training and the mazes of different length are significant. In the second half, when performance has been affected by the experience of the preceding test trials, the main differences are obscured by the significant interaction between schedules and mazes.

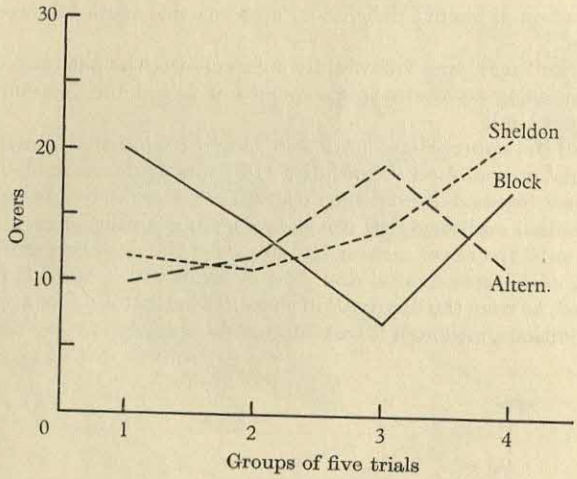


Fig. 4. Performance in the short version of the maze.

Table 1. *Analysis of variance based on number of free runs over obstacle*

(*a*) *First half* (10 *trials*)

Source	S.S.	D.F.	M.S.	*F*	*P*
Schedules	53·375	2	26·688	7·97	< 0·01
Mazes	16·333	1	16·333	4·88	< 0·05
Directions	4·083	1	4·083	1·22	N.S.
Sexes	0·750	1	0·750	< 1	N.S.
(Schedules × mazes)	(8·042)	(2)	(4·021)	(1·20)	N.S.
(Residual)	(132·667)	(40)	(3·316)	—	—
Residual (including schedules × mazes)	140·709	42	3·350	—	—
Total	215·250	47	—	—	—

(*b*) *Second half* (10 *trials*)

Source	S.S.	D.F.	M.S.	*F*	*P*
Schedules	11·375	2	5·688	—	N.S.
Mazes	6·020	1	6·020	—	N.S.
Directions	1·020	1	1·020	—	N.S.
Sexes	7·520	1	7·520	—	N.S.
Schedules × mazes	29·292	2	14·646	4·52	< 0·05
Residual	129·585	40	3·240	—	—
Total	184·812	47	—	—	—

Discussion

The results obtained must be regarded as substantially in agreement with prediction. The long maze provides an unambiguous final common path, and thus the more crucial test, but shows clearly the progression from alternation, to Sheldon and to

block schedules which should result from the suggested application of Deutsch's postulates. The Sheldon (1964) experiment therefore appears to present no difficulty for Deutsch's theory.

Even in the early test trials the block group did not choose consistently in the 'over' direction. This is partly explained by the residual 'under' tendency shown in Fig. 1 (*c*) to result from contiguity, so that a finite probability of an 'under' response should remain despite any greater number of reinforced 'over' trials than were given here. Apart from the number of training trials given, the absolute level of performance should be affected by three factors. (*a*) The strength of the presumably innate 'under' preference existing at the inception of the experiment: Deutsch (1960) suggested that the chains of links 'can be regarded as having been set up either through learning or through genetic factors.' (*b*) The value of the increments to learning arising from each 'over' or 'under' traverse of the maze: one of the unfortunate results of the qualitative nature of Deutsch's system is that the degree of reduction of resistance between contiguous links, the amount of 'inhibition' generated by failure of a sequence and the effect upon facilitation of the excitation concurrently present, are all unknown. (*c*) The effect of any tendency to alternate, owing to the temporary refractoriness of links which follows the firing of an attached analyser: evidence provided by Still (1966) suggests that Deutsch's postulates do not perhaps adequately explain the phenomena of alternation; it is nevertheless an empirical fact that substantial alternation occurs with reinforcement of both paths to a common goal box (Sutherland, 1957).

It should be noted that, whatever the level of responding reached as a result of forced-choice training, the behaviour of all three groups during free trials should tend gradually to converge at a mid-point. Given, for example, a block-reinforced animal in whom path O is strongly connected, each of the few 'under' responses to be expected will confer a relatively great advantage upon path U owing to the amount of available excitation. 'Over' performance should therefore deteriorate relatively quickly. Conversely, in an alternated animal with strong U preference the second of any pair of 'over' responses will have a large effect, while further U responses will presumably reach an asymptote, so that the resistance of the two paths should gradually equalize. These effects will of course be subject to wide fluctuations depending upon the actual sequence of responses which are made.

The form of explanation proposed extends the application of Deutsch's system, which was primarily designed to account for place learning under positive reinforcement, to an example of behaviour in which reinforcement is conditional upon response. In order to deal with conditional behaviour in a Skinner box, Deutsch (1960) introduced a different explanation, proposing as an extra postulate that an inhibitory connexion is set up between a conditional signal and the goal response. The postulation of this kind of mechanism seems to constitute an unsatisfactory explanation, inasmuch as it appears somewhat arbitrary and articulates poorly with the rest of the system. Equally, the analysis applied to the present experiment appears strained in application to the Skinner box problem, since pressing a lever rather than going direct to the food trough constitutes a detour of the kind the system is designed to eliminate. It would however be of interest to compare block with alternation procedures in the acquisition of lever-pressing, since choice of the lever detour would

be predicted if the pattern of responding made for lower resistance towards the lever than towards the food trough.

The predictions for the short maze of the present experiment are difficult to evaluate. The final sections of path may be regarded as in common or in parallel, since it is not clear to which goal-box cues the analysers correspond. Further, the animals spend some time in the goal box, with many opportunities to reverse their position. Since Pritchatt & Holding (1966) have shown that reversed connexion of links may mediate learning, it appears likely that the neural map of the final section of path is of considerable complexity. Speculation concerning the effects shown in the later test trials on the short maze therefore seems unprofitable, while results for the early test trials were of course in agreement with prediction.

References

Deutsch, J. A. (1960). *The Structural Basis of Behaviour*. London: Cambridge University Press.

Pritchatt, D. & Holding, D. H. (1966). Guiding Deutsch's model in reverse. *Br. J. Psychol.* **57**, 17–23.

Sheldon, M. (1964). Response learning. *Nature, Lond.* **202**, 1141–2.

Still, A. W. (1966). Spontaneous alternation and pattern of reinforcement. *Quart. J. exp. Psychol.* **18**, 109–18.

Sutherland, N. S. (1957). Spontaneous alternation and stimulus avoidance. *J. comp. physiol. Psychol.* **50**, 358–62.

(*Manuscript received* 17 *August* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 93–100
Printed in Great Britain

BASAL CONDUCTANCE LEVEL AND MOTOR PERFORMANCE

By R. S. CORTEEN
Psychology Department, University of Edinburgh

The relations between log. basal conductance and performance on three tasks were investigated. Significant relations were found with end spurt and reminiscence on the pursuit rotor, with overall performance on a dotting task, and with abnormally slow responses and optimum response speed in reaction time. These results are consistent with a view of basal conductance as a measure of 'tonic' activation or the ergotropic–trophotropic dimension of Hess.

Basal conductance level has long been regarded as a measure of 'alertness' (Thouless, 1925), or more recently, activation (Duffy, 1962). It is necessary to distinguish between the measurement of basal conductance and the measurement of short-term changes in basal conductance which occur as the result of discrete stimulation (G.S.R.). There is considerable confusion between the two and this has led to both being regarded as measures of general long-term activation. In a recent paper Corteen & Blackman (1965) indicated that the G.S.R. is probably a peripheral index of activity in a short term, or 'phasic' activation system and that it is, at best, only indirectly related to long-term, or 'tonic' activation. It is assumed that basal conductance is a measure of this 'tonic' activation which suggests that basal conductance might be related to performance in situations which require prolonged attention or vigilance, as distinct from performance which requires relatively brief attention to environmental detail.

Specifically the experiments to be described below were based upon an *a priori* hypothesis that basal conductance level is a reasonable index of Hess's (1964) ergotropic–trophotropic dimension. The extreme ergotropic state is characterized, according to Hess, by an attempt at high efficiency combined with reasonable expectation of success while the extreme trophotropic state is 'experienced as bodily and mental tiredness with a disinclination to make any effort at all'. It is not difficult to see a close resemblance between these states and what one would expect to be characteristic states of extreme levels of tonic activation, although it would appear reasonable to extend the activation dimension to the levels of sleep. Hess explicitly avoids drawing a parallel between this dimension and activity in the 'opposing' sympathetic–parasympathetic systems, though he admits that the similarities are great. This implies that basal conductance may not be a perfect measure of this dimension, mediated as it almost certainly is by sympathetic 'tone', (Wang, 1957, 1958). The parallel, however, ought to be sufficient for reasonable relations to be observed.

Despite a vast literature there have been surprisingly few reported findings of any relation between basal conductance level and performance, and those that do exist tend to be ambiguous or barely significant. Freeman (1940) reported a curvilinear inverted-U relation between skin resistance and reaction time. This was apparently confirmed by Schlosberg (1954), but Schlosberg & Kling (1959) failed to repeat these findings and concluded that the relation was unreliable. Other reported relations are equally nebulous.

The reason for this lack of definite findings could either be a basic flaw in the hypotheses suggesting the relation or, more probably, inadequate instrumentation. In spite of the rather depressing history of this area of experimentation, it was decided to embark upon a series of experiments designed to explore the relation between basal conductance and motor performance, in the belief that use of a highly sensitive apparatus to measure conductance might reveal relations which had hitherto remained obscure.

The experiments all took the form of inter-individual correlational studies. This was decided upon despite the arguments of Malmo (1959) in favour of intra- rather than inter-individual studies. The essential artificiality involved in attempts to manipulate the environmental situation in order to induce changes in activation was an initial consideration. In the event, it appears that tonic activation is not particularly susceptible to environmental manipulation and the only alternative in intra-individual studies is to take measurements at many different times in order to obtain a range of values. This procedure, apart from being very time consuming, would have created difficulties in taking realistic concurrent measurements of performance.

The actual relations studied were between log. basal conductance and (i) pursuit rotor performance, (ii) simple reaction time, (iii) performance on a dotting task. The expectation was that there would be positive relations between log basal conductance level and performance levels.

METHOD

Experiment I

Subjects

Twenty-one male subjects were used. All were first-year students from the psychology department of the university who had no experience of a pursuit rotor and who had not previously had their skin conductance measured. Their ages ranged from 19 to 27, with a mean of 21·5 yr.

Apparatus

The apparatus used to measure basal skin conductance was as described by Corteen & Blackman (1965). The pursuit rotor consisted of a gramophone motor and turntable, mounted in a wooden box and geared to run at 60 rev/min. A black plastic disk, 12 in. in diameter, was mounted on the turntable and a brass disk, 0·5 in. in diameter, was set into the plastic disk, the centre of the brass disk being 4·5 in. from the centre of the motor spindle. The wand used by the subject consisted of a 10 in. metal rod with a 4 in. plastic handle covering one end. From the other end hung a 3 in. long tightly wound spring, covered in insulating material, with a brass stud, $\frac{1}{8}$ in. in diameter, attached to its lower end. The rotor and wand were so wired that when the end of the wand and the brass disk came into contact a circuit was completed which started a clock. As soon as contact was broken the clock stopped. The clock could readily be read to an accuracy of 0·01 sec. An accurate stop-watch was also used.

Procedure

Skin conductance was measured before and after each full block of trials on the pursuit rotor. This was done while a short list of words was read to the subjects. These conductance measures were averaged to give a mean conductance level during the trials. Unfortunately, attempts to measure conductance during the trials proved abortive. The electrodes constantly became displaced in the middle of trials, resulting in recording pen breakages and cessation of the trials. All attempts to make the palmar electrodes more secure ended in failure.

Pursuit rotor performance was measured during three sets of five trials each, giving fifteen trials in all. The procedure for each subject was as follows. Before the first set of trials subjects were shown what they had to try to do, namely hold the end of the wand on the brass disk while the disk was rotating. They were told that each trial would be followed by a short break, during which the disk would continue to rotate, and that at the end of five trials they would be given a

5 min rest before the next series of trials. They were then given the wand and the motor was switched on. One second before they were due to start they were given a 'ready' signal and then told to begin. The stop watch was started at the moment the word 'begin' was uttered. A trial lasted for 30 sec, the words 'stop now' being uttered when the 30 sec period was completed. There then followed a 30 sec break, during which the subject remained standing in front of the still rotating disk. During this period the experimenter noted the time on target and reset the recording clock. This continued for five trials. At the end of the fifth trial the rotor was switched off and the subject was offered a chair. The rest period lasted for 5 min, during which experimenter and subject indulged in general conversation. Questions about the nature and aims of the experiment, which were almost inevitable, were evaded at this stage. Subjects were given an account of the experiment only when the session was completed. At the end of the 5 min break the procedure outlined above was repeated for another five trials and again, after a second 5 min break, for a further five trials.

Apart from actually seeing the end of the tracking wand in contact with the target disk, the subject could also confirm that contact had been successfully made by listening for the sound of the recording clock which made a readily audible sound when running. There is evidence to show that auditory confirmation of successful target tracking produces superior performance more readily than when no such confirmation is present (Reynolds & Adams, 1953).

Experiment II

Subjects

Twenty-seven subjects were used, twelve male and fifteen female. All subjects were from the second year Psychology class and their ages ranged from 19 to 37, though only two subjects were older than 23 yr. Eleven subjects had had skin conductance measures taken before but all were unfamiliar with the dotting task.

Apparatus

The skin conductance recorder was as before. The dotting task was evolved from an instrument ascribed to McDougall which was designed to measure hand–eye co-ordination. The basis of the apparatus actually used was an old 4-channel E.E.G. paper drive. This was mounted at an angle of 20° to the horizontal in a wooden box with the metal paper runner almost flush with the top of the box. The box was covered with a lid which had a slot $3\frac{1}{2}$ in. wide and $\frac{3}{8}$ in. deep cut in it to expose a part of the paper runner. The runner was wired to a 12 V battery in circuit with an impulse counter and the metal stylus. When the stylus was touched to the metal paper runner the circuit was completed and the counter registered the contact. In operation the paper drive was fed with E.E.G. paper with $\frac{1}{8}$ in. diameter holes punched in it. The aim of the task was to hit as many holes with the stylus as possible. Each hit was registered on the counter; misses were not registered. On trial runs this task seemed to be one which demanded constant attention and was challenging enough to grip the interest of the subjects. The paper used consisted of a length of E.E.G. paper 4 in. wide and twenty 12 in. segments long, plus a blank lead-in segment. The first three segments were perforated with 180 holes, $\frac{1}{8}$ in. diameter, sixty holes per segment, randomly arranged. This was followed by a blank, unperforated segment, which was followed by three perforated segments, which were again followed by a blank segment, and so on. Each length of paper consisted, therefore, of five sets of 180 perforations spread over three segments each, with a blank segment coming between each set. Thus each length of paper, when running, presented to the subject five trials with a short interval between each trial. The length of trial and interval depended upon the speed of the driving motor. The speed selected for this experiment was 1 in./sec; in fact the motor ran rather more slowly and each trial lasted 40 sec with a 13 sec interval between trials.

Procedure

Basal skin conductance was measured before and after the rapid dotting trials.

For the rapid dotting task the subjects were instructed as follows: 'This task is to study the speed of your hand–eye co-ordination. A short time after I start the motor you will begin to see holes passing beneath this slot, I want you, with this stylus [handed to subject], to try to hit as

many of these holes as possible. You will not be able to hit them all but try to hit as many as possible. When you hit a hole you will hear a sharp click. That will be the counter recording a hit. I will be listening and watching for double hits, so do not try to hit one hole more than once. There will be five trials in each set with a short break between each trial. At the end of the first set of trials there will be a break of five minutes. This will be followed by a second set of trials, another break, and a third and final set of trials. Is that clear? Now get ready to begin.'

The trials proceeded in the way indicated in the instructions. It was possible to record double hits by listening for a characteristic double click of the impulse counter and an appropriate correction was made to the score for each trial. During the 5 min rest-break, subjects were given a card sorting task while the experimenter loaded the paper for the next set of trials.

Experiment III

Subjects

Twenty-six subjects took part in the experiment, seven male and nineteen female. All subjects were from the second year psychology class. No subjects had had skin conductance measures taken before but all had taken part in an ordinary class reaction-time experiment using similar apparatus to that described below.

Apparatus

The skin conductance recorder was as before. The reaction-time unit was mounted on a 3 ft. by 2 ft. table, divided by a 2 ft. high partition which ran the width of the table. On the experimenter's side was a 'decatron' timer which could read from 0 to 0·9999 sec and a control panel containing a foreperiod time-selection switch with positions marked from 1 through to 6, and an activating switch. On the subject's side there was a red neon bulb mounted on the partition at eye level, and a reaction switch.

The apparatus worked in the following way. The experimenter selected the desired foreperiod (the period between warning buzzer and reaction stimulus) with the foreperiod time-selection switch. This operated silently so that the subject could not count the clicks and estimate the coming foreperiod. Once the foreperiod had been selected the experimenter turned the activating switch to the 'on' position. This simultaneously activated a buzzer and started the timing operation. When the foreperiod had run its course, controlled by a C–R time-constant circuit of the appropriate value to produce the desired foreperiod, the 'decatron' timer started simultaneously with the stimulus light coming on. The subject then pressed his reaction switch which stopped the timer and turned off the stimulus light. As soon as he had done this the experimenter switched the activating key to the 'off' position. This, again simultaneously, sounded a buzzer and locked the timer. If the subject released his switch before the second buzzer (i.e. with the circuit still active) the timer would continue to record and the genuine reaction time would be lost.

Reaction times and conductance levels were recorded on a printed form, on which a random sequence of foreperiods had been entered before each experimental session by entering a list of sixty numbers between 1 and 6 which were as random as possible, allowing for the fact that ten entries of each value had to be made.

Procedure

Subjects were asked to wash their hands. When this had been done they were seated in front of the reaction time apparatus and the electrodes of the skin conductance recorder were attached. The bridge of the recorder was then balanced and left at low sensitivity until the start of the experiment.

Subjects were then told: 'This is an experiment to measure your speed of reaction to a light. It is similar to the experiment you have done in class. After I have finished these instructions, I shall say nothing more to you. You will simply hear a buzzer and at a certain, variable time after the buzzer the right-hand light in front of you will come on. When it comes on I want you to press down the right hand key as quickly as possible. Hold it down until you hear the second buzzer. You should sit with your hand resting lightly on the key at all times. The series of trials will be quite long; it will last, in all, about forty-five minutes.'

Results

In the pursuit rotor experiment (Expt. I) it was found that absolute scores did not relate significantly to basal conductance level. It was found, however, that two other indices of performance did relate significantly to log. basal conductance; they were (*a*) mean reminiscence between sets 1 and 2, and 2 and 3, and (*b*) and change (improvement from trial 4 to 5 in sets 1 and 2). The results are summarized in Table 1 (*a*). Their theoretical significance is discussed below, together with other experimental results.

Table 1. *Log. basal conductance in relation* (*a*) *to reminiscence and end change measures of pursuit-rotor performance* (*see text*), (*b*) *to performance in the rapid dotting task*

(*a*)

Subject	Log. basal conductance	Reminiscence (mean)	End change (mean)
1	2·02	+0·8	+2·0
2	1·87	+0·1	+1·3
3	2·16	+1·1	+3·0
4	2·10	+0·1	+0·7
5	2·35	+0·7	−0·3
6	1·77	+4·0	+1·0
7	2·16	+2·6	−2·0
8	1·43	+2·8	+1·2
9	1·90	+4·7	+2·4
10	1·68	+6·8	−4·9
11	1·49	+15·0	−1·0
12	1·48	+3·1	−3·6
13	1·63	+0·2	−2·7
14	1·52	+1·8	−3·9
15	2·27	+2·3	+4·7
16	1·53	+2·5	+0·9
17	2·33	+2·3	+2·0
18	1·49	+7·2	−1·6
19	2·28	−2·0	+3·0
20	1·79	+3·4	+1·8
21	1·75	+4·3	−4·6
Mean	1·86	+3·04	−0·6
S.D.	0·31	3·44	2·7

$r_{LBC/REM} = -0{\cdot}50$ ($P < 0{\cdot}05$).
$r_{LBC/EC} = +0{\cdot}54$ ($P < 0{\cdot}01$).

(*b*)

Subject	Log. basal conductance	Set 1	Set 2	Set 3
1	2·20	49·2	64·0	61·6
2	2·20	77·4	81·8	82·4
3	1·88	33·6	46·6	45·0
4	2·12	37·8	43·2	45·4
5	1·68	36·6	45·6	61·4
6	1·63	50·8	48·9	51·0
7	1·95	50·2	56·0	56·0
8	1·63	49·8	52·0	53·2
9	1·68	47·6	48·2	50·2
10	1·82	43·0	44·2	54·6
11	2·01	58·2	62·4	61·6
12	1·68	31·6	41·0	42·2
13	1·89	50·2	50·2	53·6
14	1·68	72·6	65·8	59·8
15	1·95	59·8	62·0	70·6
16	1·91	48·8	61·0	59·0
17	1·92	76·0	88·8	87·8
18	1·56	38·0	39·6	46·8
19	1·48	28·4	33·4	43·6
20	1·98	49·2	54·0	56·0
21	1·94	78·0	76·8	73·8
22	2·08	36·4	57·2	61·4
23	1·61	34·8	38·2	49·3
24	1·78	50·8	59·4	55·0
25	1·98	51·2	65·2	70·0
Mean	1·84	49·6	55·4	58·0
S.D.	0·20	14·0	13·4	11·4

$r_{LBC/S.1} = +0{\cdot}38$ ($P < 0{\cdot}1$).
$r_{LBC/S.2} = +0{\cdot}59$ ($P < 0{\cdot}01$).
$r_{LBC/S.3} = +0{\cdot}53$ ($P < 0{\cdot}01$).

In the dotting task experiment (Expt. II), relations were observed between absolute performance and basal conductance. Mean performances over all five trials for each set were calculated and correlated with log. basal conductance. The results are summarized in Table 1 (*b*).

In the reaction time experiment, (Expt. III), where basal conductance was measured concurrently with performance, mean reaction times for sets of ten trials were correlated with log. basal conductance measured during these particular trials.

This gave a set of six correlations, for three of which $P < 0{\cdot}05$ and for two of which $0{\cdot}05 < P < 0{\cdot}10$. In addition, a set of correlations was calculated after abnormally slow reaction times had been eliminated (arbitrarily defined as reaction times more than 1·5 S.D. above the mean). Finally, correlations between conductance and the five fastest reaction times in a set were calculated, giving a further set of six correlations. The results are summarized in Table 2.

Table 2. *Product-moment correlations of log. basal conductance with mean reaction time, with mean reaction time excluding abnormally slow reactions, and with the mean of the five fastest reactions*

	Mean RT	Mean RT (excluding abnormally slow)	Mean five fastest RT
Set 1	−0·37 ($P < 0{\cdot}10$)	−0·38 ($P < 0{\cdot}10$)	−0·47 ($P < 0{\cdot}05$)
Set 2	−0·23 (N.S.)	−0·19 (N.S.)	−0·37 ($P < 0{\cdot}10$)
Set 3	−0·44 ($P < 0{\cdot}05$)	−0·38 ($P < 0{\cdot}10$)	−0·35 ($P < 0{\cdot}10$)
Set 4	−0·45 ($P < 0{\cdot}05$)	−0·26 (N.S.)	−0·30 (N.S.)
Set 5	−0·37 ($P < 0{\cdot}10$)	−0·27 (N.S.)	−0·24 (N.S.)
Set 6	−0·43 ($P < 0{\cdot}05$)	−0·48 ($P < 0{\cdot}02$)	−0·44 ($P < 0{\cdot}05$)

DISCUSSION

In general the results point to a significant relation between basal conductance and a range of motor performances. None of the observed relations was significantly non-linear, which is of considerable interest in view of the consistently hypothesized inverted-U relation between activation and performance. It may be, of course, that the range of activation levels sampled was too narrow to show up any curvilinearity, but there is an alternative argument which would suggest that there is only a simple linear relation between tonic activation level and performance on tasks requiring continuous sensory input, and that curvilinear relations reported in other studies are due to an interaction of more than one 'intensive' dimension.

The theoretical arguments supporting the above suggestion are too lengthy to be presented in full, but they involve the hypothesis that 'drive' in a context can be expressed as the discrepancy between the optimum stimulation level for a given level of tonic activation and the actual level presented by a situation. The greater the discrepancy the greater the 'drive' to increase or reduce the stimulation to the optimum level. If it is assumed that the greater the drive the lower the efficiency, in the sense that any device hunting to achieve a predetermined setting will increase in efficiency the closer it is to that setting (Pribram, 1961), then the efficiency of performance is not a simple direct function of activation level but a function of activation level and the level of stimulus input determined by the task in a particular context. The implications are then that high activation level would never normally result in poor performance on complex (i.e. high stimulus input) tasks, but that high drive levels, which might be confused observationally with high activation, would. The fact that 'drive', as rather unusually defined here, has been confused with activation is not surprising in view of the vague and general definitions of activation which have become current.

One reason for the confusion might be in the conception that tonic activation can be readily manipulated by changes in the environment. In other words, activation level can be altered by increasing or decreasing the level of direct and indirect stimu-

lation, thus affecting performance. The suggestion put forward here is that it is not tonic activation which is altered by this manipulation, but 'drive'. If we accept tonic activation as being equivalent to Hess's ergotropic–trophotropic dimension, then it is highly unlikely that direct stimulation would have any influence upon it, at least in the short term, though it might be directly manipulable by the administration of drugs. This conception must further distinguish tonic from phasic activation, which is directly dependent upon stimulation. The introduction of a concept of drive, dependent upon but not identical with activation, which is inversely related to efficiency, might also serve to explain many results in experiments relating indices of 'anxiety' or 'neuroticism' to performance.

Of the results of the experiments outlined above, a comparison of the relations observed between basal conductance and pursuit-rotor performance and between basal conductance and performance on the dotting task is of the greatest interest. For the pursuit-rotor conductance was not related to the absolute level of performance. High conductance or activation subjects tended to maintain or improve performance up to the last trial in a set and to show little reminiscence, while low activation subjects showed the reverse tendency. For the dotting task the results were quite different. Neither intra-set improvement nor reminiscence was related to conductance, but the overall performance level was positively related.

The explanation of this contrast probably lies in the different nature of the two tasks. The pursuit rotor situation involves a considerable amount of learning and there tends to be continuous improvement throughout the sets of trials. This effect is much less marked in the dotting task where, typically, there is very little improvement after the first two trials. It is possible, therefore, that in the pursuit rotor situation high activation subjects, coping more efficiently because of their lower 'drive' level, tend to develop fewer inappropriate responses within a set of trials than low activation subjects. They will thus show superior performance at the end of a set but less reminiscence, since there are fewer inappropriate responses to die out during the rest period. This explanation implies a differential forgetting theory of reminiscence (McGeoch & Irion, 1952). In the dotting task, which involves much less complex sensori-motor co-ordination, inappropriate responses are much less likely to occur and so the rest period will provide no relative advantage to the low activation subjects. Differences in activation, therefore, show up as differences in absolute performance on this task.

The findings in the reaction-time experiment seem to show two different effects: (*a*) that conductance level is closely related to the number of abnormally slow responses—(the higher the conductance level the fewer the slow responses, see Table 3); (*b*) that conductance level is also related to the optimum response times (the higher the conductance level the faster the mean optimum response time). This implies that there is a group of response times for all subjects, neither abnormally slow nor tending towards the optimum, which will be unrelated to conductance. This was, in fact, found to be true. Correlations between mean intermediate response times and conductance level were all totally non-significant.

Of the two factors found to be related to conductance, the first, percentage of abnormal responses, is not surprising. If activation is related to general efficiency in tasks requiring long-term attention or vigilance, then more 'blocks' are to be

expected at lower levels of activation. The second finding, however, is of some interest in view of Taylor's (1965) suggestion that optimum reaction times might not be related to activation level. Taylor used a different criterion of optimum reaction time, the fastest time in a series, while in this study the mean of the five fastest times was used. This was done in order to minimize the effect of any anticipatory responses but recalculation of the correlations using only the fastest times in any series of ten did not alter the coefficients to any significant extent. It would, therefore, appear that the greater efficiency characteristic of high activation subjects is reflected in their optimum speed of reaction as well as in their lower rate of 'blocking'.

Table 3. *Abnormally slow responses as a function of log. basal conductance*

($\chi^2 = 10{\cdot}13$; D.F. 4; $P < 0{\cdot}05$.)

Log. basal conductance	1·2–1·49	1·5–1·69	1·7–1·89	1·9–2·09	2·1–2·39
% abnormal RT	11·8	9·6	9·3	2·3	3·0

The generally low correlations observed in the reaction-time study probably reflect the fact that phasic activation must play a part in speed of reaction. As phasic activation is only indirectly related to tonic activation, this probable relation must diminish the degree of direct correspondence between reaction time and tonic activation.

In conclusion it may be said that the experiments described above illustrate the validity of basal conductance as a behaviourally relevant measure and that the types of relation observed are not inconsistent with a conception of basal conductance as an index of tonic activation or the ergotropic–trophotropic dimension. There is a distinct lack of evidence for any inverted-U relations and it is possible that these will occur only when 'drive' rather than activation, is manipulated in intra-individual studies.

REFERENCES

CORTEEN, R. S. & BLACKMAN, A. R. (1965). Skin conductance change and sensory sensitivity. *Br. J. Psychol.* **56**, 431–7.

DUFFY, E. (1962). *Activation and Behavior*. New York: Wiley.

FREEMAN, G. L. (1940). The relationship between performance level and bodily activity level. *J. exp. Psychol.* **26**, 602–8.

HESS, W. R. (1964). *The Biology of Mind*. Chicago: University of Chicago Press.

MCGEOCH, J. A. & IRION, A. L. (1952). *The Psychology of Human Learning*. London: Longmans, Green.

MALMO, R. B. (1959). Activation: a neurophysiological dimension. *Psychol. Rev.* **66**, 367–86.

PRIBRAM, K. H. (1961). Implications for systematic studies of behaviour. In *Electrical Stimulation of the Brain*, ed. D. E. Sheer, pp. 563–74. Austin: University of Texas Press.

REYNOLDS, B. & ADAMS, J. A. (1953). Motor performance as a function of click reinforcement. *J. exp. Psychol.* **45**, 315–20.

SCHLOSBERG, H. (1954). Three dimensions of emotion. *Psychol. Rev.* **61**, 81–8.

SCHLOSBERG, H. & KLING, J. W. (1959). The relationship between 'tension' and efficiency. *Percept. mot. Skills* **9**, 395–7.

TAYLOR, D. H. (1965). Two dimensions for reaction time distributions. *Nature, Lond.* **206**, 219–20.

THOULESS, R. H. (1925). The causes of the continuous change of resistance observed in psychogalvanic experiments. *Br. J. Psychol.* **16**, 5–15.

WANG, G. H. (1957). The galvanic skin reflex I. *Am. J. phys. Med.* **36**, 295–320.

WANG, G. H. (1958). The galvanic skin reflex II. *Am. J. phys. Med.* **37**, 25–57.

(*Manuscript received* 25 *March* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 101–104
Printed in Great Britain

SENSORY PROCESSING AND THE RECOGNITION OF FORMS IN NURSERY-SCHOOL CHILDREN

By BRUCE DENNER and SHELDON CASHDAN
Indiana University

Recognition of forms under varying sensory input conditions was studied in three groups of pre-school children. One group inspected an object visually while the two other groups were given the additional experience of handling the object either directly or encased in a ball of clear plastic (haptic and manual conditions). Recognition and verbal recall measures were obtained two days later. No differences in verbal recall were observed. In agreement with previous research, performance by the 'visual only' group was worse than by the other two groups. Since the recognition by first two groups was equally good, it is suggested that the component underlying haptic facilitation of learning may simply be the amount of concrete activity involved.

There are a number of references in the British literature (O'Connor & Hermelin, 1963; Birch & Lefford, 1963) and numerous statements in the Russian literature (Razran, 1961; Zaporozhets, 1965; *Psychological Science in the USSR*, 1961) which suggest that whereas early in ontogenesis the 'contact sensory modes' are the dominant means of acquiring information about the world, later in ontogenesis there is a shift to the 'distant sensory modes'. Much of the work in this area is difficult to interpret and contradictory findings have been reported, but it appears that young children (2–5 yr) not only prefer to touch and manipulate objects, but they also retain information acquired by these means more adequately.

Two distinct explanations are possible. On the one hand, the various sensory modalities may develop unevenly; presumably the contact modes of tactual perception (textural contact) and haptic perception (manipulative contact) mature prior to the distant modes of vision and audition. As a consequence very young children are more efficient learners when they are permitted to make tactual contact with stimulus materials and to manipulate them (Tarakanov & Zinchenko, 1960). An ordinal development of the sensory systems receives further support from studies which purport to show that developmentally more primitive subjects, e.g. autistic children (O'Connor & Hermelin, 1963) and retarded children (Hermelin & O'Connor, 1961), prefer tactile stimuli and recognize objects more effectively when they are allowed to inspect them manually. But, contrary to the foregoing generalization, other studies suggest that haptic perception (exploratory handling of objects) matures later than visual perception (Page, 1959; Zinchenko & Lomov, 1960). Hence something other than the ontogenetic priority of one class of sensory modes over another appears to be involved.

Alternatively, a sequence of formal stages, like those of Werner (1948) and Piaget (1952), may describe the development of information processing independently of sensory system. From this point of view, the child who first organizes his world in terms of concrete, overt activity, comes to organize his world by more contemplative means, i.e. perception and thought. One corollary of this proposition that has been borne out in research is that very young children tend to respond on the basis of the intensity rather than the meaning of a stimulus (Luria, 1961). But perhaps more

critical is the work reported by Gibson (1963). For both visual and tactual perception it appears that the amount of overt *activity* involved in processing information decreases during development; there is a decrement in the amount of eye-movement involved in visual inspection and a decrement in the amount of manual activity ingredient in haptic perception. The present research represents a first attempt to clarify some of the problems posed by earlier work. Starting from an experimental situation (Tarakanov & Zinchenko, 1960) in which it was reported that nursery-school children recall objects better if they explore them manually, procedures were introduced to determine whether it was the activity with the object or the haptic exploration which facilitated recognition.

Three groups of nursery-age children were selected for the study. One group inspected a geometric form visually and haptically, another inspected the same form visually but was not permitted to touch it, while the third group inspected the form visually but it was encased in a clear plastic ball which they were allowed to handle.

If it is the activity rather than the actual physical contact with the form which is critical for acquiring information, no difference would be expected between the group who explored the form haptically and the group who manipulate the ball containing the form. Hence it is necessary to include a 'visual only' condition to control for the possibility that a null finding might reflect only an effect of visual dominance in the other two conditions. If performance is worse in the 'visual only' group than in the other two groups, it follows that it is the activity which facilitates the processing of information rather than haptic exploration or visual search alone.

METHOD

Subjects. Three groups of eight normal nursery-school children ranging in age from 3 to 4 years were equated for age and I.Q. (Peabody Picture Vocabulary Test). These values appear in Table 1. The subjects were randomly assigned to one of the three treatment conditions: visual inspection, visual plus haptic inspection, and visual plus manual (in the ball) inspection.

Materials and procedure. The materials consisted of two sets of the same geometric forms adopted from the Seguin Form Board. The forms were a hexagon, cross, square, and triangle whose outside dimensions were such that they could be snugly encased in a transparent sphere. The stimuli were cut from hard-board ⅛ in. thick and painted bright red. Each of the forms was encased in a clear plastic ball 3·75 in. in diameter. The inspection stimulus figure in all cases was the hexagon.

During the inspection period only the stimulus figure (hexagon) was present. The child was first told that he was going to play a 'remembering game' in which he was going to be shown something he had to remember. The child's attention was then drawn to the figure and its name was pronounced clearly (the word 'hex' was used). The child was then required to repeat the name and told to look at the 'hex' for a while (30 sec) or to handle it, depending on the condition. After this period the child was required to repeat the name. If he remembered the name he was given a trinket. If he failed, the experimenter repeated the name and asked the child again. The child would then repeat the name and was given a trinket. Finally, the child was told that the experimenters would return in 2 days and that if he remembered what had been shown to him he would receive another trinket. The child was made to paraphrase the final instructions in order to ensure his comprehension of the task.

In the test condition (recognition–recall), the forms rested approximately 1 in. apart on a wooden board. The figures were kept out of the subject's view by a metal cover placed over the board until the instructions had been given. During the test for recognition the stimuli were placed on the board in the same position for all subjects, and 'hex' was second from the subject's left. While the cover was still over the materials, the child was asked if he remembered what had

been shown to him the other day. He was told that it was underneath the cover but that there were other things there also. He was then asked to point to the object previously shown to him. While the cover was being removed the experimenter cautioned the subject against an impulsive choice by requiring him to look at all the figures before he responded. After the child had pointed to the form he was asked if he remembered its name; he was given another trinket irrespective of whether his response was correct or not.

Results

The percentage of correct form-recognition responses and the percentage of correct naming responses in the three groups are presented in Table 1. Whereas all children in groups 1 (visual and haptic inspection) and group 2 (visual and manual inspection of figure in ball) correctly pointed out the stimulus figure, only 63% of the children in group 3 (visual inspection) correctly pointed to the figure. In this last group two children pointed to the triangle and one pointed to the cross. The hypergeometric distribution gives an exact probability of $P < 0{\cdot}027$ for this pattern of results (Feller, 1957). Hence, the reliability of difference between group 3 (visual (inspection) and the other groups is better than $P < 0{\cdot}05$).

Table 1. *Mean age, mean* I.Q. *and percentages of correct recognition and naming responses within the three treatment conditions*

Group	Treatment condition	Mean age (months)†	Mean I.Q.†	Correct recognition responses (%)*	Correct naming responses (%)†
1	Visual + haptic	44·5	108	100	25
2	Visual + manual (stimulus in transparent globe)	47·1	111	100	50
3	Visual	44·6	108	63	38

* $P < 0{\cdot}05$. † Differences not significant.

Table 1 also shows the percentage of children in each group who correctly named the figure. In most cases the children responded with some name, but only a clearly articulated statement of the name 'hex' was scored as correct. The exact probability for this distribution of results is $P < 0{\cdot}085$; hence the different treatments did not lead to reliably different recollection of the names.

Discussion

The finding that young nursery-school children are more likely to recognize a form if they manipulate it than if they merely inspect it visually is in general agreement with the literature. But the finding that manipulation without haptic involvement combined with visual exploration of an object is as efficacious as haptic-visual exploration sheds some light on the development of sensory processing. These results suggest that it is not the visual experience alone or the haptic contact with objects which is of ultimate importance to the young child; haptic involvement with an object may facilitate perceptual recognition simply because of the activity intrinsic to this mode of processing information. Hence, the two points of view concerning the development of sensory modes can be reconciled: (*a*) the maturation of contact modes

prior to the distant modes with (*b*) the developmental decrement in concrete activity involved in the processing of information. One may assume that in development there is a shift from the use of sensory means which maximize the degree of overt concrete activity to the use of sensory means that minimize the degree of activity.

The findings with regard to the recall of the name of the form are by contrast somewhat difficult to interpret. This may be due to inadequacies of the system that was used for scoring the responses. In point of fact there were two classes of children; children who did not respond (approximately 50 % in each group), and children who responded with some approximation to the word. A more meaningful analysis of these responses is contingent upon a workable classification of the verbalizations. It seems clear, however, that overt verbalization of a name at the time of recognition is not essential for correct recognition to occur. This result is puzzling since verbal mediation would have been expected to play an important part. All the children had, of course, been trained to use the name, and, in the last analysis, it is impossible to know what part may have been played by the inner-speech of the silent child.

In sum, it appears that active manipulation is one of the important means by which the young child comes to know objects. The role of verbal processes remains to be clarified.

REFERENCES

BIRCH, H. D. & LEFFORD, A. (1963). Intersensory development in children. *Monogr. Soc. Res. Child Dev.* **28**, no. 5, serial no. 89.

FELLER, W. (1957). *An Introduction to Probability Theory and its Applications*, vol. 1. New York: Wiley.

GIBSON, E. J. (1963). Perceptual development. In *Child Psychology* (62nd yearbook of the National Society for the Study of Education), ed. H. W. Stevenson, chapter 4. Chicago: University of Chicago Press.

HERMELIN, B. & O'CONNOR, N. (1961). Recognition of shapes by normal and subnormal children. *Brit. J. Psychol.* **52**, 281–4.

LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour.* New York: Liveright Press.

O'CONNOR, N. & HERMELIN, B. (1963). Sensory dominance in autistic children and subnormal controls. *Percep. mot. Skills*, **16**, 920.

PAGE, E. (1959). Haptic perception: A consideration of one of the investigations of Piaget and Inhelder. *Educ. Rev.* **2**, 115–24.

PIAGET, J. (1952). *The Origins of Intelligence in Children.* New York: International Universities Press.

Psychological Science in the USSR (1961). Washington: U.S. Joint Publications Research Service.

RAZRAN, G. (1961). The observable unconscious and the inferable conscious in current Soviet psychophysiology. *Psychol. Rev.* **68**, 127–8.

WERNER, H. (1948). *Comparative Psychology of Mental Development.* New York: International Universities Press.

TARAKANOV, V. V. & ZINCHENKO, V. P. (1960). Comparative analysis of touch and vision. Communication VI: Voluntary and involuntary learning of form at pre-school age. *Doklady Akad. Ped. Nauk RSFSR* **5**, 41–52.

ZAPOROZHETS, A. V. (1965). The development of perception in the preschool child. In *European Research in Cognitive Development*, ed. P. H. Mussen, chapter 7. *Monogr. Soc. Res. Child Dev.* **30**, no. 2, serial no. 100.

ZINCHENKO, V. P. & LOMOV, B. F. (1960). The functions of hand and eye movements in the process of perception. *Problems of Psychology* **1**, 12–26.

(*Manuscript received* 19 *October* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 105–110
Printed in Great Britain

CONDITIONED ORIENTING REACTIONS IN TWO-MONTH-OLD INFANTS

By JAROSLAV KOCH
Institute for the Care of Mother and Child, Prague-Padolí, Czechoslovakia

Visual orientation is one of the earliest and most basic responses of the young infant to the human adult (Ambrose, 1961; Brackbill, 1958; Kagan & Lewis, 1965; Wolff, 1963). At first the child fixates the adult, and later there are visual pursuit movements that follow a moving person and locate him among other objects. The human adult is the strongest stimulus influencing the behaviour of the child and its whole mental development. Whether this is so from the first day of life or whether the relationship develops as a result of repeated contacts between child and adult in biologically significant situations is a matter for investigation. One special aspect of this complex problem was selected for study: the development of conditioned head rotation in infants in an experimental situation. The aim was to study the conditions of reinforcement of orienting responses. For this purpose different kinds of reinforcer were used, the development of conditioned orienting responses was recorded, and conclusions were reached about the strength of each reinforcer.

Method

Subjects

The subjects were fifty normal healthy infants at the end of their second month of life. They lived with their mothers in the Institute for the Care of Mother and Child in Prague-Podolí.

Apparatus

During the experiment the child was placed in an enclosed crib (100 × 60 × 60 cm) with the head resting on a simple device for recording turning movements. On each side of the child's head there was a window (30 × 30 cm and 50 × 50 cm) where the reinforcing stimulus could appear. The windows could be closed and opened by curtains. Near the feet of the child was the source of sound stimuli.

Procedure

A conditioned orienting head rotation to a sound was established by presenting one of five types of reinforcer: (1) face and voice of the mother; (2) face and voice of a stranger; (3) various noise-making toys changing from trial to trial; (4) strange adult showing a different toy on each trial; (5) strange adult taking the child out of the crib after each response, rocking and coddling it and performing other activities that please children.

After the presentation of the sound that was used as the conditioning stimulus, three different types of response were recorded:

(1) *Conditioned response*: turning the head within 8 sec to the side on which the reinforcer appeared.

(2) *Primary response*: turning the head only after the 8th second when the side-window was illuminated and the face of an adult calling the child, or a toy emitting a sound, appeared.

(3) *Zero reaction*: not turning the head even to the reinforcing stimulus. When this occurred the experimenter turned the child's head to attain reinforcement.

There were ten sessions of the same type for each child; in each session there were ten paired presentations of the condition stimulus and the reinforcing stimulus. The sessions lasted 10–12 min and the children had only one session a day.

RESULTS

The results obtained in experiments with fifty infants at the end of their second month of life are reported. The results of other experiments with 3-month-old infants are referred to briefly for comparison.

Figs. 1–5 show changes in the percentage of the conditioned orienting responses (full-line curves), and changes in the percentage of both primary and conditioned responses (dashed-line curves). Each figure presents the result of ten infants in ten sessions.

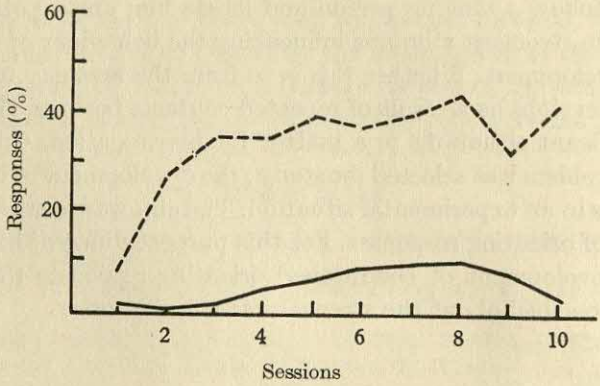


Fig. 1. Conditioning curves for orienting head rotation reinforced by presentation of the face and the voice of the mother in 2-month-old infants. ———, conditioned responses; – – –, primary and conditioned responses.

Fig. 1 shows that it is difficult to establish a conditioned orienting reaction in 2-month-old infants when orienting responses are reinforced by the appearance of the face and the voice of the mother in the side-window (reinforcement type 1). In the seventh and eighth sessions only 10 % of the conditioned stimuli were followed by a conditioned reaction and in the two sessions that follow, the percentage of conditioned responses had a declining tendency. In some infants the disappearance of the conditioned and also of the primary orienting responses was quite remarkable. These children reacted in the eighth and tenth session in a negativistic way and did not turn their heads toward the mother or even turned their heads away from her. They actively resisted the attempt of the experimeter to turn their heads toward the mother and their mood deteriorated with restlessness and crying. Evidently the children were oversatiated with the reinforcer. Similar results were obtained in 3-month-old children in whom the orienting head turning was reinforced in the same way, by the appearance of the mother. In the third session more than 40 % of the conditioned stimuli were followed by a conditioned response, but in later sessions the percentage of these reactions also declined.

Fig. 2 shows results from ten infants in whom the orientating response of head-turning was reinforced by the appearance of a stranger speaking gently (reinforce-

ment type 2). The percentage of conditional orienting responses steadily increased and in the tenth session 20 % of the conditioned stimuli evoked a conditioned response. The percentage of conditioned responses with this type of reinforcement increased faster than in experiments in which the reactions were reinforced by the appearance of the mother. Similarly, in 3-month-old infants the percentage of conditioned orienting responses reinforced by a stranger rose faster than when reinforced by the mother.

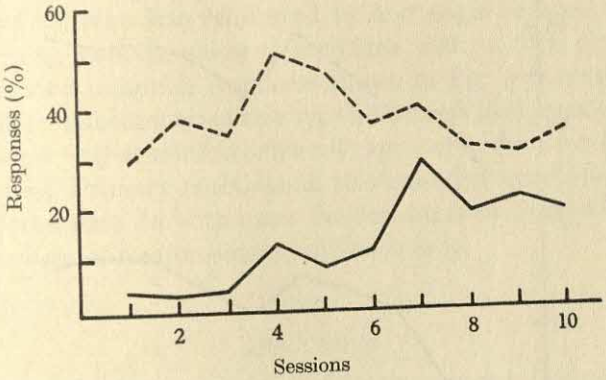


Fig. 2. Conditioning curves for orienting head rotation reinforced by the presentation of the face and the voice of a stranger in 2-month-old infants. ———, conditioned responses; – – –, primary and conditioned responses.

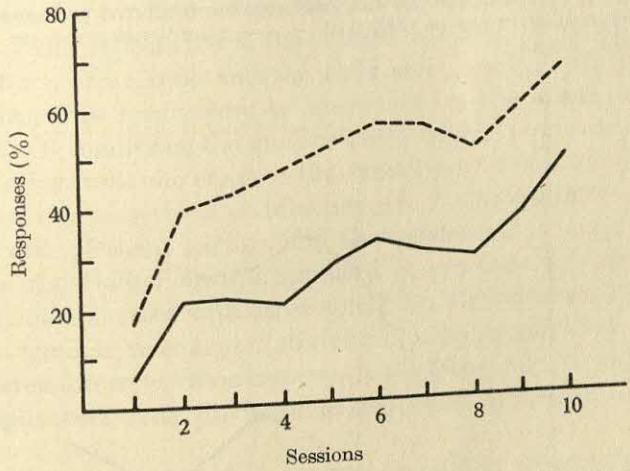


Fig. 3. Conditioning curves for the orienting head rotation reinforced by the presentation of various toys changing from trial to trial. ———, conditioned responses; – – –, primary and conditioned responses.

Fig. 3 indicates the change of orienting responses reinforced by showing a different toy on each trial (reinforcement type 3). Conditioned responses appeared earlier with this type of reinforcement than with reinforcement by the appearance of the same person on each trial. A steady increase occurred without a later decline. In the tenth session the children reacted positively to 40 % of the conditioned stimuli. In 3-month-

old children in experiments with reinforcement of this type the percentage of conditioned responses rose faster than with reinforcements types 1 or 2. An attempt was also made to reinforce orienting reactions using the same toy on each trial and under these conditions the conditioned orienting responses developed with the same speed as when always reinforced by the appearance of the same stranger.

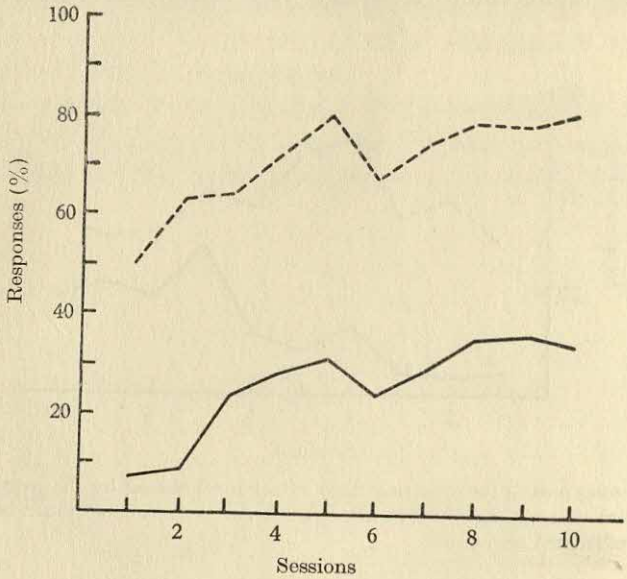


Fig. 4. Conditioning curves for the orienting head rotation reinforced by the presentation of a stranger showing a different toy on each trial. ———, conditioned responses; – – –, primary and conditioned responses.

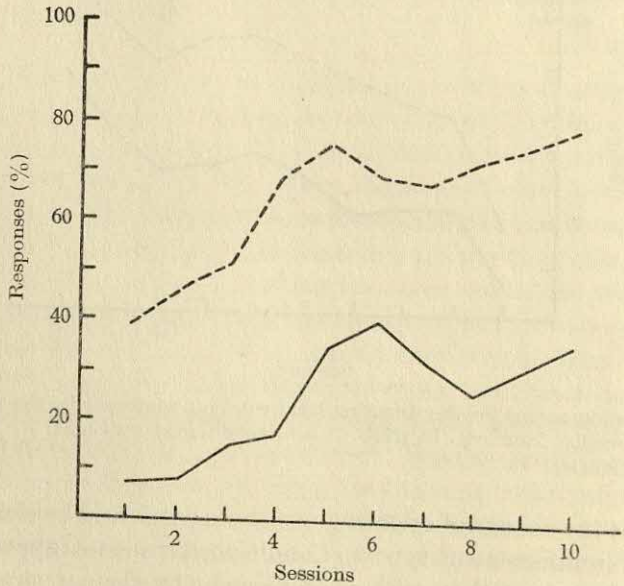


Fig. 5. Conditioning curves for the orienting head rotation reinforced by the presentation of a stranger rocking and coddling the infant. ———, Conditioned responses; – – –, primary and conditioned responses.

In the other groups of 2-month-old children conditioned orienting reactions were reinforced by more complex stimuli. Fig. 4 shows the change in percentage of orienting reactions reinforced by a stranger who showed different toys to the child. The changes that occurred were nearly identical with the changes produced by reinforcement with the toys alone (Fig. 3). A summation effect of both components (stranger + different toys) was not present. But the percentage of the primary responses rose significantly in comparison with primary responses reinforced only by different toys.

Much the same results were obtained from the final group of ten infants in whom the orienting head rotation was reinforced by a stranger rocking and coddling the child and performing various pleasant activities with it. The change in the percentage of conditioned orienting responses shown in Fig. 5 is much the same as in Figs. 3 and 4. While reinforcement of this type influenced the formation of conditioned responses in the same way as reinforcement of type 3 or 4, there was again a difference in primary reactions. Primary reactions in this case developed in the same way as with type 4 reinforcement. In both cases the development of primary reactions was faster than in the case of reinforcement with toys only.

Discussion

The aim was to find out how conditioned orienting responses develop when different kinds of reinforcer are used. From the results we may draw direct conclusions about the properties of the orienting reflex; the properties of social relationships can only be inferred indirectly and with the greatest caution.

The establishment of conditioned reactions in experiments in which head rotation was reinforced by the appearance of the mother was very difficult. These reactions extinguished after a short time and the child soon became satiated with the reinforcing stimulus. This result may be attributed to the unnatural experimental situation. In natural conditions the mother keeps appearing in other situations: she feeds the child, afterwards she changes the swaddling-clothes, she bathes the child, carries him, shows him something, coddles him, etc. These situations are for the most part connected with pleasant feelings and the satisfaction of needs. In our experimental situation the mother always appeared in precisely the same way and her appearance was not connected with other activities. The mother in our experimental arrangement is a familiar, unchanging stimulus. Such a stimulus had an effect only in first sessions as a reinforcer, but soon its power to elicit the orienting reflex extinguished and further application even produced oversatiation revealing itself in aversive behaviour.

The stranger also appeared in the same way from trial to trial, but he was a new and unknown object for the child. The fact that the stranger appeared only in the experimental situation caused him to have more reinforcing strength than the mother, notwithstanding that he also appeared in the same way from one trial to another. Experiments with 3-month-old infants in whom the orienting reaction was always reinforced with the same toy show that the percentage of the conditioned orienting reactions increased in the same manner as when reinforced by a stranger. Both of these reinforcers have the common property that they are new, although they are repeated.

The pairing of orienting reactions with differing toys increased significantly the percentage of conditioned orienting reactions. Each toy in the experimental situation was something new that never appeared during the session again.

With more complex reinforcers—for example, a stranger showing different toys or a stranger rocking the child in his arms—the percentage of conditioned responses changed in much the same way as when reinforced by showing different toys without the appearance of an adult. But the percentage of primary responses reinforced by the complex stimulus increased faster in comparison with the increase of primary responses reinforced by different toys only. While the complex stimuli did not have a summative effect on conditioned orienting reactions in 2-month-old infants, a summative effect was revealed in primary reactions. In experiments in which we used complex reinforcing stimuli, the child increased primary head turning to the toy or to the person and afterwards it was able on the basis of the improved primary reactions to improve upon the conditioned head turning.

CONCLUSIONS

1. If a stimulus, with which the 2-month-old infant is acquainted in other situations, is used in the same manner and in the same situation over and over again, then this stimulus weakens as a reinforcer for the establishment of a conditioned orienting response. Eventually the orienting reaction extinguishes, and the stimulus begins to elicit aversive behaviour.

2. If a new, hitherto unknown stimulus appears in a certain situation repeatedly, it has a stronger reinforcing effect in establishing a conditioned orienting response than a stimulus that has been experienced in other situations. But even a new stimulus with repetition loses its reinforcing strength, and the orienting reaction formerly elicited by this stimulus extinguishes.

3. The strongest reinforcer for the establishment of conditioned orienting responses is a new and changing stimulus.

4. The simultaneous action of more than one stimulus evoking the orienting reaction (the action of compound) stimuli does not accelerate the establishment of conditioned orienting responses in 2-month-old infants, but it does accelerate the development of primary orienting reactions. The development and strengthening of these primary reactions is the precondition of the establishment of conditioned orienting responses.

REFERENCES

AMBROSE, J. A. (1961). The development of the smiling response in early infancy. In *Determinants of Infant Behaviour*, ed. B. M. Foss, pp. 179–201. London: Methuen.

BRACKBILL, Y. (1958). Extinction of the smiling response in infants as a function of reinforcement schedule. *Child. Devel.* **29**, 115–224.

KAGAN, J. & LEWIS, M. (1965). Studies of attention in the human infant *Merrill-Palmer Quart.* **11**, 95–122.

WOLFF, P. H. (1963). Observation on the early development of smiling. In *Determinants of Infant Behaviour*, ed. B. M. Foss, vol. II, pp. 113–67. London: Methuen.

(*Manuscript received* 2 *July* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 111–126
Printed in Great Britain

RIGIDITY, DRIVE AND CONDITIONING IN NEUROTICS*

By R. W. PAYNE
Eastern Pennsylvania Psychiatric Institute, Temple Medical School, Philadelphia

C. T. NETLEY
Institute of Psychiatry, London

and R. B. SLOANE
Temple Medical School, Philadelphia

Four of Brengelmann's questionnaire measures of rigidity, Brengelmann's Drive Scale, the Taylor Manifest Anxiety Scale, the Maudsley Personality Inventory and the Mill Hill Vocabulary Scale, were given to thirty neurotic and thirty normal female subjects. In the first of two conditioning experiments, the eyeblink response was conditioned to a tone, an air puff being the UCS. In the second, the finger flexion reaction to an unavoidable electric shock was conditioned to a tone. GSR and heart-rate responses were measured simultaneously.

Brengelmann's previous findings that the rigidity questionnaire scores were significantly correlated with drive, anxiety and neuroticism questionnaire measure were confirmed. However, the predicted relations between rigidity and four of Malmo's physiological measures of drive (arousal) were not found. Although conditioning was clearly obtained, Brengelmann and Field's finding that eyeblink conditionability and questionnaire rigidity are related was not verified, nor did the conditioned finger response correlate with these rigidity measures. The data suggested that, for the GSR and heart-rate responses, only 'pseudo-conditioning' was obtained. The possible implications of this last finding are discussed.

Although many studies have been carried out, it has still to be demonstrated that there is any such general personality trait as 'rigidity'. In 1915, Lankes administered a questionnaire and a wide range of tests of 'perserveration' to forty-seven normal subjects, and reported that all the measures given correlated as expected with the questionnaire. A more recent study by Cattell, Dubin & Saunders (1954) also reported a general factor of rigidity and suggestibility. However, a large number of other studies using mainly performance and perceptual tests all failed to obtain a general factor of rigidity (Jasper, 1931; Shevach, 1937; Notcutt, 1943; Kleemeier & Dudek, 1950; Scheier & Ferguson, 1952; Goodstein, 1953; Rim, 1953; Pitcher & Stacey, 1954; Scheier, 1954; Forster, Vinacke & Digman, 1955; Davids, 1955; Wolpert, 1955). French (1955) and Moffitt and Stagner (1956) have suggested that a factor of rigidity might be demonstrable if the tests are given under stress, but their results and those of Applezweig (1954) lend little support to this hypothesis.

There is some evidence, however, for more restricted rigidity factors. Cattell (1946), Cattell & Tiner (1949) and Schaie (1955) found that motor tests of rigidity tended to produce two relatively specific and independent factors.

While there is no evidence for a generalized trait of rigidity embracing motor, perceptual and mental tasks, one specific type of mental rigidity, 'Einstellung' rigidity (Duncker, 1945), has been clearly defined and frequently studied. This is usually defined operationally in terms of Luchins' 'Water-Jar' test, in which subjects are required to solve problems of measuring specified quantities of water using only

* This study was supported by Canadian Federal–Provincial Mental Health Grant no. 605-5-318.

three jars of known capacity (Luchins, 1942, 1951; Luchins & Luchins, 1950). Luchins has demonstrated that, after a specific solution pattern has been consistently found to work (reinforced), other simple types of solutions are more diffcult to perceive (the Einstellung effect). Cowen, Wiener & Hess (1953) found that the Water-Jar test correlated significantly with a very similar letter maze test of Einstellung rigidity, and Guilford, Frick, Christensen & Merrifield (1957) found it to be related to some other tests of logical reasoning. Nevertheless, Einstellung rigidity is probably a fairly specific aspect of behaviour, as Levitt & Zuckerman (1959) did not find it to be related to other performance measures of rigidity.

There have been a number of attempts to relate Einstellung rigidity to other aspects of behaviour. Correlations with rigidity questionnaires have produced conflicting results (Cowen & Thompson, 1951; Goodstein, 1953; Pitcher & Stacey, 1954; Schmidt, Fonda & Wesley, 1954; Rokeach, McGovney & Denny, 1955; Wolpert, 1955). Similarly, attempts to demonstrate that prejudiced or ethnocentric people are especially rigid in this sense have been inconclusive. (Rokeach, 1948; Levitt & Zelen, 1953; Brown, 1953; Jackson, Messick & Solley, 1957).

One line of research has, however, produced consistent results. When normal subjects are given the Einstellung test and some perceptual measures of rigidity under stress, they become more rigid. (Cowen, 1952*a*, *b*; Brown, 1953; Pally, 1954; Smock, 1955*a*, *b*; Erikson & Wechsler, 1955; Moffitt & Stagner, 1956; Maher, 1957). Payne (1960) has suggested that Einstellung rigidity is a type of discrimination learning in which only the reinforced responses are acquired and the non-reinforced responses are extinguished. Therefore, learned performance under high drive states (e.g. anxiety) could be expected to be better (or more 'rigid' in this context). One study suggests that, perhaps for this reason, neurotic patients are more rigid than normal people (Pervin, 1960).

A second line of research which has tended to yield consistent results, is that involving some specific questionnaire measures of rigidity. Wesley (1953) developed a rigidity questionnaire which has proved to be reliable, and which was shown to correlate both with independent clinical ratings of rigidity and with a 'shift' score on a card sorting concept formation test (cf. Zelen & Levitt, 1954). Perhaps the most extensive set of studies of questionnaire measures of rigidity has been that carried out by Brengelmann and his colleagues. Brengelmann (1960*a*, *b*) has worked with five questionnaire measures of rigidity, a revision of the Nigniewitzky (1955, 1956) rigidity scale (NR scale), a revision of Gough's (1952) California psychological inventory of rigidity (CPIR scale), a revision of the Rokeach (1953) dogmatism scale (DO scale), a revision of the Nigniewitzky (1955, 1956) intolerance of ambiguity scale (IA scale), and a questionnaire measure of extreme response set (Brengelmann's revision of the Soueif, 1958, 'personal friend check list', PFCL scale). Brengelmann (1960*a*), studying eighty-eight normal and 105 neurotic and psychotic subjects, found that all five scales intercorrelated significantly and that they correlated with a questionnaire measure of 'drive' designed to be uncontaminated with anxiety or neuroticism (Brengelmann's DR scale). He also found positive correlations between these rigidity measures and the Taylor (1953) Manifest Anxiety Scale (TMAS scale), and with the neuroticism (N) scale of Eysenck's (1959) Maudsley Personality Inventory (MPI). Anxious and neurotic subjects, both as diagnosed by psychiatrists and by the

questionnaires, were rigid, a result consistent with the findings derived from the Einstellung experiments already discussed.

In addition, Brengelmann has investigated the relation between questionnaire rigidity and a set of learning measures derived from his Figure Reconstruction Test, in which subjects are required to learn a pattern of shapes presented visually over a large number of trials. In the first study (Brengelmann, Hahn, Pedley & Amato, 1960), rigid subjects proved to be poor at this particular kind of pattern learning. This result was replicated by Brengelmann & Brengelmann (1960*a*) in a study using 100 German nationals. In a further study, Brengelmann & Brengelmann (1960*b*), using slightly different learning measures, found a curvilinear relation between rigidity and learning, a moderate degree of rigidity unexpectedly being associated with the poorest learning.

More recently, Field & Brengelmann (1961) have explored the relation between questionnaire rigidity and eyeblink conditioning. Using a sample of thirty-three prisoners, they found a multiple correlation coefficient of 0·459 between speed of eyeblink conditionability and rigidity, perhaps because both are a function of high drive.

Although there have been no studies of the relation between Einstellung rigidity and these particular questionnaires, it would appear to be the case that both sorts of rigidity may be associated with high drive or states which produce high drive (neuroticism and anxiety).

The present study was designed to explore further the relation between questionnaire rigidity, drive and conditioning, in an attempt to verify and to extend the findings of Field & Brengelmann (1961) that they are all positively correlated.

Method

Subjects

Neurotic group. This consisted of thirty female patients from the psychiatric ward of the Kingston General Hospital. No subject was included who was thought to be psychotic or who was known to be suffering from organic disease. Attempts were also made to exclude subjects who were on high doses of drugs on admission. Otherwise those selected were females chosen in the order of their admission between the ages of 16 and 65 yr. A primary feature of all patients tended to be anxiety. Except for night sedation, all patients selected were kept off drugs for 3 days prior to testing.

Normal group. These were thirty normal female volunteers matched for age and intelligence with the neurotic group. The subjects consisted largely of nurses aids from the Kingston General Hospital, but included secretaries and housewives as well. All were paid for their participation in the experiment.

Psychological tests

Mill Hill Vocabulary Test (Raven, 1958). This vocabulary test was given as a control measure in order to match the two groups for general intellectual level. The I.Q. scores quoted are based on a standard transformation from the published norms (Raven, 1958), defining the mean as 100 and the standard deviation as fifteen for each age group. This test was individually administered to the subjects, along with the questionnaires described below, before the beginning of the conditioning experiment.

Brengelmann's (1960*a*) *Drive* (*DR*) *Scale.* This scale was constructed to measure drive in the sense of a need for achievement. It is a self-administered questionnaire, in which the subject must rate a number of characteristics as 'true' or 'false' as they apply to himself (e.g. 'When doing

a task, I feel that my individual reputation is at stake.'). The score is the number of items answered in the 'high-drive' direction.

Brengelmann's (1960*a*) *revision of the Nigniewitzky* (1955, 1956) *Intolerance of Ambiguity* (*IA*) *Scale*. This is the first of four of Brengelmann's rigidity questionnaires included in the present study. It is a sixteen-item 'true-false' scale including such statements as 'First impressions about people are usually correct.' The score is the number of 'intolerant' responses.

Brengelmann's (1960*a*) *revision of the Gough* (1952) *California Psychological Inventory of Rigidity* (*CPIR*) *Scale*. The second rigidity questionnaire consists of fifteen 'true-false' statements (e.g. 'Our thinking would be a lot better off if we would just forget about words like "probably", "approximately", and "perhaps".') The score is the number of items answered in the 'rigid' direction.

Brengelmann's (1960) *revision of the Rokeach* (1953) *Dogmatism* (*DO*) *Scale*. The third rigidity questionnaire consists of sixteen 'true-false' items such as 'To compromise with our political opponents is to be guilty of appeasement.' The score is the number of 'dogmatic' responses.

The Taylor (1953) *Manifest Anxiety Scale* (*TMAS*). This questionnaire was included to provide a measure of anxiety, since it had been used in previous studies.

The Maudsley Personality Inventory (*MPI*). Eysenck's (1959) measure of extraversion (E) and neuroticism (N) was included in an attempt to repeat Brengelmann's earlier results.

Conditioning experiments

In order to replicate the study by Field & Brengelmann (1961), eyeblink conditioning tests were made on all the subjects. In addition, conditioning of three other responses was studied, in order to extend the previous work: galvanic skin response (GSR), heart rate and finger withdrawal.

Half the subjects had the eyelid-conditioning and half had the finger-withdrawal experiment first, in a random order. All subjects received the psychological tests first.

The heart-rate, finger-withdrawal and GSR-conditioning experiment. The conditioned stimulus (CS) was a 1000 c/s tone generated by a 'Maico' audiometer, and delivered to the subject through an earphone to the preferred ear, the headset being balanced by a non-functional earpiece. The strength of the CS was set at about 40 db above threshold, as established by the method of limits. The subject lay with her left hand strapped to a board. Her middle finger which was free to move, rested on a steel plate mechanically linked to a spring mounted tambour so that finger movements could be recorded. The unconditioned stimulus (UCS) was an unavoidable electric shock delivered to this same middle finger through two electrodes, one attached to the proximal phalanx of the finger, and one to the distal end. The d.c. electric shock was produced by an apparatus similar to Alexander's (1958), which held the shock-current constant in spite of variations in skin resistance. A milliameter recorded the shock levels administered. GSR was measured continuously using the Fere method involving impressed current, following the general procedure of Malmo & Davis (1961). In addition, simultaneous recordings were made of heart rate using a 'Grass Model 5-P4A' pre-amplifier and tachograph.

The subject was tested in an electrostatically shielded air-conditioned room 6 ft. × 10 ft. × 6 ft. high. The experimenter and recording equipment were in an adjacent room. The experimenter could observe the subject through a one-way screen and communicate by a two-way speaker system. The subject was told the experiment involved in part a measure of her reaction to a painful stimulus. She was asked to lie down, and the various electrodes were attached. The nature of the shocking apparatus was then explained, and the subject was instructed to hold down the tambour plate with the middle finger of her left hand. The experimenter then left the room, and all further instructions were given over the speaker system. The strength of the shock (UCS) which was to be used was first determined by administering discrete shocks of 0·2 sec duration, in increasing increments of 0·1 mA until the shocks elicited a wincing response (i.e. attempt to withdraw the finger). The strength of the UCS was maintained at 0·5 mA below this level throughout the conditioning procedure. During the succeeding 5–10 min period, the subject was presented with a series of flashes of light of 0·2 sec duration from a 60 W bulb, 8 ft. from her eyes. These stimuli were distractors while the timers were set and the physiological measures returned to normal. The subject was then given instructions explaining the nature of the conditioning process. She was cautioned not to respond voluntarily, but also not to prevent her finger from moving.

A series of tones alone was then presented approximately 20 sec apart until three successive

tones failed to produce any deflection in GSR. None of the subjects made a finger withdrawal to tone alone during this period. As soon as this criterion was reached, the conditioning series proceeded without a break and without further instructions.

A trace conditioning procedure was used with a 0·2 sec tone followed 0·4 sec later by a 0·2 sec shock giving an interval of 0·6 sec between onset of CS and UCS. Fifty acquisition trials were presented with one test trial of tone alone randomly interspersed among every five acquisition trials. The conditioning trials were followed immediately by twenty extinction trials of tone alone. The presentations of the CS were separated by intervals of 30–60 sec in a random sequence.

During the extinction period, any responses which occurred within 1·5 sec of the CS were scored as conditioned responses. If three successive CSs produced no response during this period, extinction was said to have occurred, and subsequent responses were not scored. While the subjects could not avoid the shocks, it is to be noted that the strength of the shock was such that the shocks always in fact produced finger-withdrawal responses.

The GSR and heart rate were recorded simultaneously with the finger-avoidance responses. In scoring conditioned GSR responses, Becker's (1960) criteria were employed. Conditioned heart-rate responses were scored according to the criteria described by Zeaman, Deane & Wegner (1954). Conditioned heart-rate changes were scored only if the amplitude of the change (in terms of beats per minute) was half as great or greater than the amplitude of the change observed in response to CS + UCS. Again, extinction was assumed to have occurred if three successive CSs produced no GSR or heart rate responses, and no subsequent responses were scored.

The eyeblink conditioning experiment. Subjects were tested in the same room as that used for the finger-withdrawal conditioning. The subject sat in a straight backed chair with her head on a chin rest. About 3 ft. in front of her eyes, and slightly above her line of vision was a 'Lurgo' lamp with a 15 cm diameter flat white reflector. The CS consisted of an abrupt change in the level and hue (from white to reddish orange) of this reflected light source, caused by the discharge of a 2 W 'General Electric NE.40' neon lamp. The UCS was a puff of compressed air delivered at a pressure of 140 mm of mercury as measured at the point of delivery to the eye. The air puff was applied at a distance of 0·5 in. from the left eye from a 1 mm diameter jet mounted on an eye-glass frame. On the top and bottom of the left frame, two 'Grass' EEG silver disk-type electrodes were attached by flexible flat spring steel. These electrodes were covered with EKG 'Sol' conductive electrode jelly, and when the eye-glass frames were worn, the electrodes pressed firmly above and below the eyeball. The signal from these electrodes generated by movement of the eyelid was amplified and recorded on a 'Grass model 3A' EEG machine. Preliminary investigations with this recording device permitted calibration, so that it was relatively insensitive to movements of the eye, and recorded only the muscle potential generated by a gross eyelid movement. The presentation of the CS and UCS were controlled by three 'Hunter' timers. During the experiment two 15 W incandescent lamps were alight at all times to minimize beta responses (Grant & Norris, 1946).

The subject was brought into the room and informed that the experiment involved testing protective reflexes, of which an eyeblink to a puff of air is a good example. The procedure was explained, and the eye-glass frames fitted to the subject. Several trials using the UCS alone were administered in order to familiarize the subject with the procedure and to adjust the sensitivity of the recorder. The experimenter then left the test room. The subject then received a series of ten lights alone of 600 msec duration to measure the orienting responses to the CS. A conditioning series of fifty acquisition trials then followed without a break. The CS duration was 100 msec. After 500 msec, the UCS was given. The duration of the UCS was 100 msec, so that CS and UCS terminated simultaneously. Ten extinction trials of CS alone followed without a break. Throughout, the presentations of the CS were separated by intervals of 30–60 sec in a random sequence. An eyelid closure of at least 1 mm deflexion from the base-line of the polygraph recording, within the latency range of 200–525 msec after the onset of the CS, was considered to be a CR during the extinction series. The conditioning measures were scored by two judges independently; where differences occurred, they were resolved by discussion.

Results

Differences between groups on initial psychological tests

Table 1 shows the differences between the neurotic and the normal group on the tests given before the conditioning experiments were carried out. The two groups did not differ significantly in age or intelligence, suggesting that the matching was successful. As expected, the neurotics were both significantly more neurotic on the N scale of the MPI, and significantly more introverted on the E scale of the MPI than the normal group. The latter difference is not surprising as anxiety was the main feature of the neurotic group, and this is reflected by the very large significant difference on the TMAS scale. Brengelmann's drive scale did not differentiate the groups significantly, and this was consistent with his earlier results (Brengelmann, 1960*a*).

Table 1. *Differences between neurotic and normal groups (n = 30/group) in age, intelligence, neuroticism, anxiety, drive, rigidity and initial shock level*

	Normal group		Neurotic group			
	Mean	S.D.	Mean	S.D.	t	P
Age	37·7	12·8	35·7	13·1	0·60	N.S.
Intelligence (Mill Hill Vocab. I.Q.)	97·9	9·8	94·7	8·5	1·58	N.S.
Neuroticism (MPI N)	18·9	9·3	32·3	9·9	3·84	< 0·01
Extraversion (MPI E)	27·1	6·4	20·9	8·8	3·26	< 0·01
Anxiety (TMAS)	12·3	6·3	30·1	8·2	14·34	< 0·001
Drive (D)	6·5	2·6	6·3	2·6	0·29	N.S.
Rigidity (NR)	11·4	7·6	14·7	5·2	1·95	N.S.
Rigidity (CPIR)	14·9	7·0	17·0	6·5	1·24	N.S.
Rigidity (DO)	10·8	6·8	13·1	4·3	1·57	N.S.
Rigidity (IA)	16·5	8·3	21·2	5·1	2·65	< 0·01
Shock level (mA)	2·3	0·9	2·2	0·9	0·43	N.S.

Table 2. *Correlation coefficients between rigidity tests and personality tests, intelligence and age for combined groups (n = 60)*

	Rigidity scale			
	NR	CPIR	DO	IA
Drive (D)	0·25	0·444**	0·245	0·206
Anxiety (TMAS)	0·271*	0·207	0·274*	0·400**
Neuroticism (MPI N)	0·358**	0·299*	0·349**	0·447**
Extraversion (MPI E)	−0·232	−0·105	−0·015	−0·304*
Intelligence (Mill Hill Vocab.)	−0·420**	−0·429**	−0·413**	−0·503**
Age	0·225	0·193	0·193	0·138

* $P < 0·05$. ** $P < 0·01$. (2-tail tests.)

However, only one of the rigidity scales (IA) differentiated the groups, and in this respect, Brengelmann's (1960*a*) finding that neurotics were more rigid on all four scales is only partially verified. It is of interest that almost precisely the same level of shock was needed to produce the wincing response in both groups.

Relation between rigidity tests and other questionnaires

It was argued that the best way of testing the relation between the rigidity tests and the other variables was to combine the data from both the normal and the

neurotic group so as to avoid any possible restriction of the range. Accordingly, product-moment correlation coefficients were calculated between the four rigidity scales and the other personality questionnaires, as well as with age and intelligence for all sixty subjects. These results are shown in Table 2. They are more or less as expected. Rigid subjects tend to be neurotic (MPI N scale) and anxious (TMAS) and tend to have high drive (D scale). However, while all the correlations are in the expected direction and the results collectively are better than chance, not all the individual correlation coefficients reach statistical significance. Rigid subjects also tend to be introverted and less intelligent. These findings on the whole are consistent with those reported by Brengelmann (1960*a*), although they are at a lower level of statistical significance.

Table 3. *Correlations between physiological arousal measures and rigidity questionnaires for combined groups* ($n = 60$)

	Rigidity scale			
	NR	CPIR	DO	IA
Base-line skin resistance	0·057	0·106	0·113	0·160
Base-line fluctuations of skin resistance	0·158	0·112	−0·008	0·004
Base-line heart rate	0·030	−0·091	−0·115	−0·139
Heart rate change to shock	−0·049	0·001	−0·098	−0·332*

* $P < 0·05$. (2-tail test.)

Table 4. *Correlations between GSR conditioning and rigidity questionnaires for combined groups* ($n = 60$)

	Rigidity scale			
	NR	CPIR	DO	IA
No. of conditioned GSR responses (acquisition period)	0·144	0·055	−0·001	0·038
No. of conditioned GSR responses (extinction period)	0·133	0·104	0·002	0·115
Latency of CR (acquisition period)	−0·037	−0·061	−0·033	−0·013
Latency of CR (extinction period)	−0·146	−0·145	−0·160	−0·081

Arousal and rigidity

As a further test of the suggestions that rigidity might be related to drive state (Brengelmann, 1960*a*; Payne, 1960; Wesley, 1953), several physiological indices of drive or arousal were obtained. Following Malmo's (1957) suggestions, these were: (1) the base-line skin resistance in ohms; (2) fluctuations of base-line skin resistance; (3) base-line heart-rate beats/min; (4) heart-rate changes following the initial shock (in beats/min). It was recognized that none of these measures was likely to be independent of the conditioning indices used.

The correlations between these four measures of drive and the four rigidity questionnaires are shown in Table 3. Only one achieves significance at the 5% level, and this result is probably a chance finding in so large a matrix. It must therefore be concluded that the results do not support the hypothesis that questionnaire rigidity and high drive are associated.

GSR conditioning and rigidity

The product-moment correlation coefficients obtained between the rigidity questionnaires and the GSR-conditioning measures are shown in Table 4. None of these correlation coefficients is significantly greater than zero, providing no evidence for the expected relation between rigidity tests and GSR conditioning.

It might be argued that none of these rigidity scales is sufficiently reliable alone and that some combined measure would be a better criterion of rigidity. Accordingly, each subject's standard score on each rigidity test was summed to produce a single score of rigidity. The entire group was then divided at the mean into a 'high rigidity' and 'low rigidity' group of thirty subjects each. The data for the conditioning trials were combined into three blocks of three trials each, as were the data for the extinction trials, and this grouping is the basis of the trend analysis between trials reported in Table 5. As can be seen from an inspection of Table 5, contrary to expectation,

Table 5. *Trend analysis of GSR conditioning data for high and low rigid groups*

Source	S.S.	D.F.	M.S.	*F*	*P*
	Acquisition period				
Groups	0·272	1	0·272	0·138	N.S.
Error I	114·000	58	1·975	—	—
Trials	1·744	2	0·872	1·06	N.S.
Groups × trials	2·140	2	1·070	1·31	N.S.
Error II	94·701	116	0·817	—	—
Total	213·000	179	—	—	—
	Extinction period				
Groups	0·138	1	0·138	0·071	N.S.
Error I	113·078	58	1·949	—	—
Trials	33·100	2	16·550	43·609	< 0·005
Groups × trials	0·211	2	0·106	0·278	N.S.
Error II	44·022	116	0·379	—	—
Total	190·550	179	—	—	—

there are no significant between-group differences, either in the acquisition or the extinction period. It is of considerable interest, however, that in spite of the care which was taken to extinguish the initial GSR reaction to the tone alone *after* the shock levels to be used had been obtained, these data provide no evidence that any conditioning whatever was taking place. As can be seen clearly from Fig. 1, the subjects start out at a high level, and gradually produce *fewer* (but not significantly fewer) 'conditioned' GSR responses during the conditioning period. This decrease in 'conditioned' GSR responding continues (but now at a statistically significant level) throughout the extinction period (see Fig. 2). These data, therefore, strongly suggest a 'pseudo-conditioning' effect. Evidently, after the initial GSR response to the tone had been extinguished, the first shock during the conditioning period re-aroused the subjects to such an extent that a GSR startle reaction was produced by all the tones, which, if anything, extinguished (or, more correctly, 'habituated') to some extent during the 'conditioning' period, and clearly habituated during the 'extinction' period. These results may partly have been the result of the strong shock levels used,

but they throw considerable doubt on the feasibility of getting a genuinely uncontaminated measure of GSR 'conditionability' using an aversive stimulus. Accordingly the results of this part of the experiment, while technically of considerable interest, must be regarded as inconclusive so far as the hypothesis being tested is concerned.

Heart-rate conditioning and rigidity

Table 6 shows the correlations obtained between the measures of heart-rate conditioning and the four rigidity questionnaires. None of these coefficients attains statistical significance, contrary to expectation. Table 7 shows the results of the

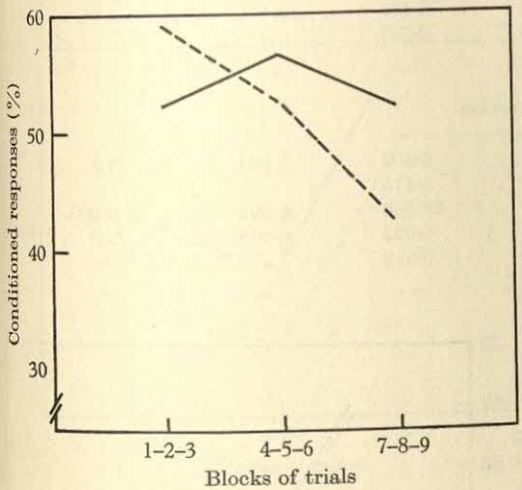


Fig. 1

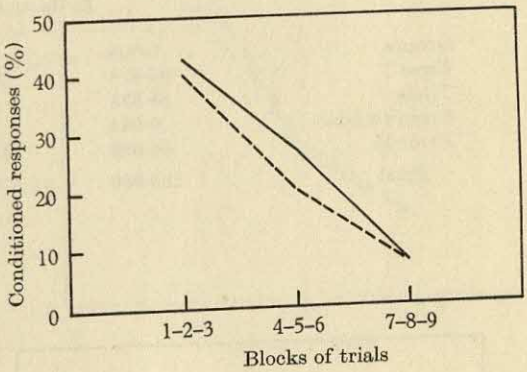


Fig. 2

Fig. 1. GSR conditioning (acquisition trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Fig. 2. GSR conditioning (extinction trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Table 6. *Correlations between heart-rate conditioning and rigidity questionnaires for combined groups* ($n = 60$)

	Rigidity scale			
	NR	CPIR	DO	IA
No. of CRs (acquisition period)	−0·149	0·026	−0·046	−0·029
No. of CRs (extinction period)	−0·055	0·183	0·103	0·082

trend analyses carried out on these conditioning measures. The data were grouped in the same way as were the GSR data. Figs. 3 and 4 show the acquisition and extinction curves for the two groups. These data are very similar to the GSR data. Again, there is no statistically significant learning curve, although this time the direction of the slope in the 'flexible' group is at least upwards during the acquisition period. The differences between the groups is, however, not significant. During the extinction period, the performance of the two groups is indistinguishable, but both show a sharp and highly significant 'extinction' curve. As with the GSR changes, these results are

probably almost entirely due to the 'sensitization' to all stimuli resulting from the shocks, since reactivity tends to *start* at a maximum level, and to drop off sharply when the shocks cease.

Table 7. *Trend analysis of heart-rate conditioning data for high and low rigid groups*

Source	S.S.	D.F.	M.S.	*F*	*P*
	Acquisition period				
Groups	0·556	1	0·555	0·347	N.S.
Error I	92·689	58	1·598	—	—
Trials	0·878	2	0·439	0·576	N.S.
Groups × trials	0·878	2	0·439	0·576	N.S.
Error II	88·244	116	0·761	—	—
Total	183·244	179	—	—	—
	Extinction period				
Groups	0·005	1	0·005	0·004	N.S.
Error I	87·878	58	1·515	—	—
Trials	58·533	2	29·266	56·498	< 0·005
Groups × trials	0·044	2	0·022	0·043	N.S.
Error II	60·088	116	0·518	—	—
Total	206·550	179	—	—	—

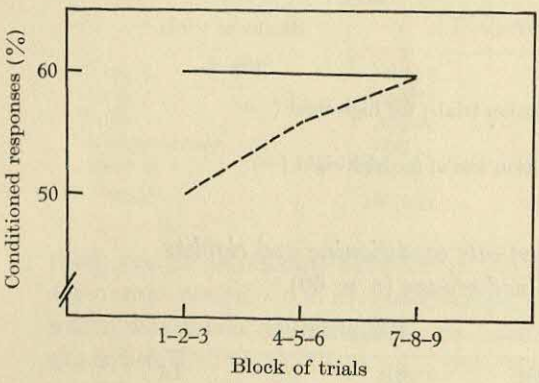


Fig. 3

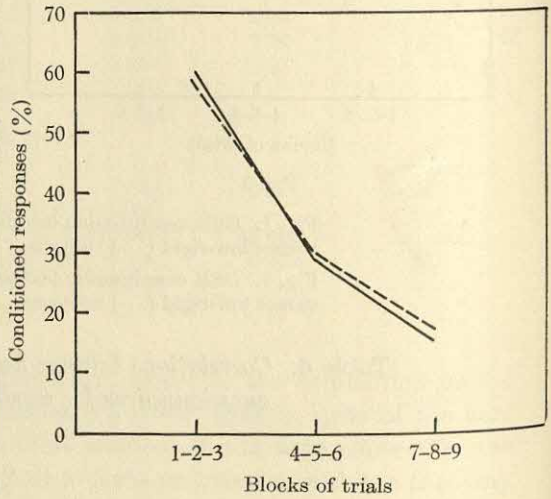


Fig. 4

Fig. 3. Heart-rate conditioning (acquisition trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Fig. 4. Heart-rate conditioning (extinction trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Finger-withdrawal conditioning and rigidity

Table 8 shows the correlations obtained between the measures of finger-withdrawal conditioning and the four rigidity measures. Again, none reaches statistical significance. Table 9 shows the results of the trend analysis of this aspect of the data. It was carried out in the same way as for the variables described above. Fig. 5 shows the

acquisition curves, and Fig. 6 the extinction curves for the two groups for this response. As can be seen, once again the two groups are not significantly differentiated. For this response, however, while there appears to be a moderately high level of

Table 8. *Correlations between finger withdrawal conditioning and rigidity questionnaires for combined groups* ($n = 60$)

	Rigidity scale			
	NR	CPIR	DO	IA
No. of CRs (acquisition period)	0·133	−0·041	0·131	−0·068
No. of CRs (extinction period)	0·080	−0·066	0·088	−0·080

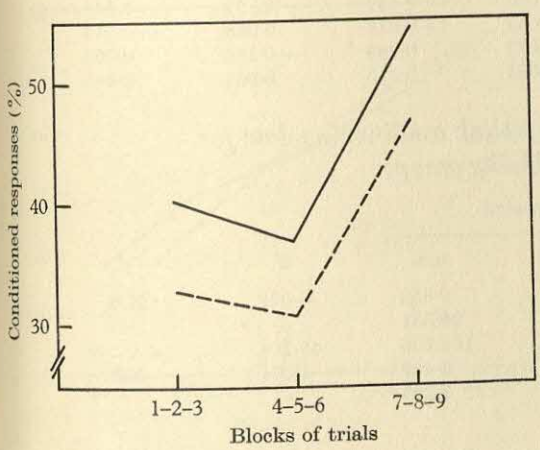


Fig. 5

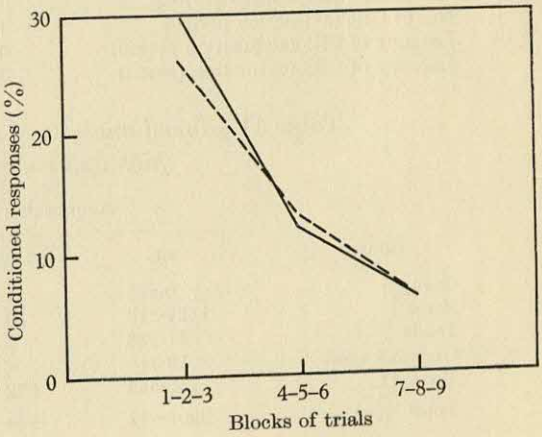


Fig. 6

Fig. 5. Finger-withdrawal learning (acquisition trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Fig. 6. Finger-withdrawal learning (extinction trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Table 9. *Trend analysis of finger-withdrawal conditioning data for high and low rigidity groups*

Source	S.S.	D.F.	M.S.	*F*	*P*
Acquisition period					
Groups	1·800	1	1·800	1·339	N.S.
Error I	77·977	58	1·344	—	—
Trials	10·077	2	5·039	6·024	< 0·005
Groups × trials	0·233	2	0·116	0·139	N.S.
Error II	97·022	116	0·836	—	—
Total	187·111	179	—	—	—
Extinction period					
Groups	0·050	1	0·050	0·038	N.S.
Error I	74·678	58	1·287	—	—
Trials	13·878	2	6·938	20·178	< 0·005
Groups × trials	0·233	2	0·116	0·339	N.S.
Error II	39·889	116	0·343	—	—
Total	128·873	179	—	—	—

initial responsiveness which could be due to sensitization, this appears to extinguish fairly quickly, and then there appears to be evidence of genuine conditioning. Unlike the GSR and heart-rate responses, the changes during the conditioning period for finger withdrawal are statistically significant. There is also a highly significant extinction effect during the extinction period.

Table 10. *Correlations between eyeblink conditioning and rigidity questionnaires for combined groups* ($n = 60$)

	Rigidity scale			
	NR	CPIR	DO	IA
No. of CRs (acquisition period)	0·092	0·137	0·152	0·148
No. of CRs (extinction period)	−0·048	−0·028	0·058	−0·019
Latency of CRs (acquisition period)	−0·072	0·008	−0·192	0·068
Latency of CRs (extinction period)	0·068	0·115	0·062	0·187

Table 11. *Trend analysis of eyeblink conditioning data for high and low rigidity groups*

Source	S.S.	D.F.	M.S.	F	P
	Acquisition period				
Groups	0·853	1	0·853	0·032	N.S.
Error I	1527·213	56	26·331	—	—
Trials	617·233	4	154·308	48·204	< 0·005
Groups × trials	13·713	4	3·428	1·071	N.S.
Error II	742·653	232	3·201	—	—
Total	2901·667	299	—	—	—
	Extinction period				
Groups	0·050	1	0·050	0·026	N.S.
Error I	110·944	58	1·912	—	—
Trials	51·078	2	25·538	53·711	< 0·005
Groups × trials	0·433	2	0·216	0·456	N.S.
Error II	55·155	116	0·475	—	—
Total	217·661	179	—	—	—

Eyeblink conditioning and rigidity

Table 10 shows the correlations obtained between the measures of eyeblink conditioning and the rigidity tests. Again none of the correlations is significant. Table 11 shows the results of the trend analysis of these data, grouped in a similar way to the other variables. Fig. 7 shows the acquisition curves and Fig. 8 the extinction curves for the two groups for this response. Again, the performance of the two groups is nearly identical, there being no suggestion of any statistically significant differences, contrary to Field & Brengelmann (1961). For this response, unlike the others, there is virtually no suggestion of an initial sensitization effect. At the beginning of the conditioning period, the subjects give virtually no conditioned responses, and there is a progressive and highly significant increase in conditioned responses during the acquisition period. Again there is a highly significant extinction effect during the extinction period, and there is even a suggestion of the classically shaped extinction curve.

Discussion

The present study on the whole verified Brengelmann's (1960*a*) finding that rigid subjects, as assessed by his questionnaires, also tend to give anxious, neurotic and introverted responses to other questionnaires, and to show evidence of a high level of drive as assessed by a questionnaire. The evidence was less clear that patients who are diagnosed as neurotic tend to show more questionnaire rigidity than normal people, although the differences were all in the direction found by Brengelmann. However, when four physiological measures of drive were correlated with the rigidity questionnaires, the results were statistically non-significant. These findings suggest that,

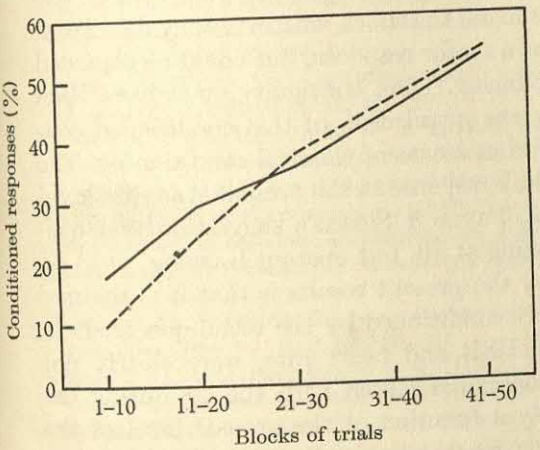


Fig. 7

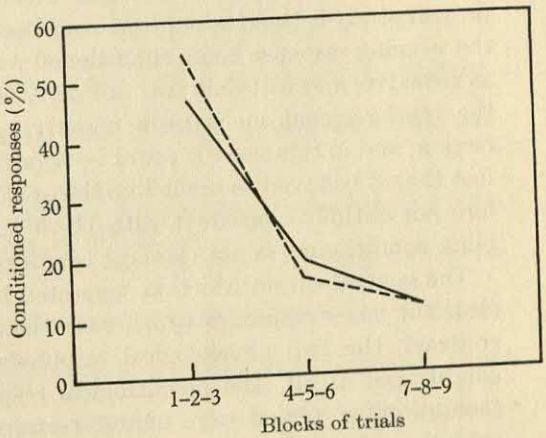


Fig. 8

Fig. 7. Eyeblink conditioning (acquisition trials) for high-rigid (—) versus low-rigid (- - -) subjects.

Fig. 8. Eyeblink conditioning (extinction trials) for high-rigid (—) versus low-rigid (- - -) subjects.

although a common questionnaire factor tends to underly the performance of the subjects on all these measures, there is no evidence that it extends to physiological responsiveness. It may, therefore, be dangerous to identify questionnaire with physiological 'drive'. Consistent with these findings was the failure to replicate the finding of Field & Brengelmann (1961) that questionnaire rigidity and eyeblink conditioning are positively correlated.

An interesting aspect of these data is the shape of the learning and extinction curves. Eysenck (1965) and his colleagues have frequently used measures of both eyeblink conditioning and GSR conditioning when testing the theory that conditionability in general is associated (through the cortical inhibition/excitation balance) with extraversion/introversion. Nevertheless, the 'two-factor' theory maintains that classical conditioning and (operant) learning are independent phenomena. GSR conditioning is a good example of 'classical conditioning' as the response itself is reflexive, and it is difficult to conceive of it as an 'operant'. Emitting a GSR response is not especially pleasurable; often the reverse. It would appear, therefore, that contiguity is the only relevant condition for reinforcing GSR conditioning. The eye blink response

on the other hand is ambiguous. Like GSR, it could be regarded as a reflexive response. However, it is a motor response, and is usually an operant in the sense that it reduces drive by removing an irritant. Thus eyeblink conditioning could conceivably be regarded as an example of learning by virtue of drive reduction. Davidson, Payne & Sloane (1964) have produced some evidence that the acquisition of the finger-withdrawal response, when it can be made to avoid an electric shock, and is thus clearly operant, is completely independent of GSR (classical) conditioning. They go on to argue that if eyeblink conditioning is also really operant learning, it is not surprising that there is so little clear-cut evidence for significant correlations between both GSR and eyeblink conditionability, and introversion (Eysenck, 1965).

The special interest of the finger-withdrawal conditioning procedure used in the present study is that the subjects could not avoid the shock whatever they did. Thus, the wincing response being conditioned was a motor response, but could be regarded as reflexive, and certainly was not drive reducing. Thus, contiguity must have been the reinforcement mechanism underlying the acquisition of this conditioned connexion, and in this sense it could be regarded as a case of classical conditioning. The fact that it behaved so much like the eyeblink response in the present study is therefore not entirely consistent with Davidson, Payne & Sloane's suggestion that eyeblink conditioning is not classical conditioning at all, but operant learning.

The generalization which *is* suggested by the present results is that it is the two clear-cut *motor* responses which were clearly conditioned by the technique used. In contrast, the two physiological responses, GSR and heart rate, were clearly not conditioned at all. The physiological responses associated with the CS during the 'conditioning' period were almost certainly a function of the arousal level of the subjects, who were evidently sensitized to react to any discrete stimulus when aroused, and were thus evidence only of 'pseudo-conditioning.' These clear-cut results suggest that, even under less aversive experimental conditions (strong shocks were used in the present study), GSR and heart rate 'conditioning' scores may be in part a measure of physiological arousal level, and may thus have little to do with cortical excitation/inhibition balance. This would make the use of this type of conditioning measure of limited value in testing this particular personality theory.

References

Alexander, L. (1958). Apparatus and method for the study of conditioned reflexes in man. *A.M.A. Archs. Neurol. Psychiat.* **80**, 629–49.

Applezweig, D. G. (1954). Some determinants of behavioral rigidity. *J. abnorm. soc. Psychol.* **49**, 224–8.

Becker, W. C. (1960). Cortical inhibition and extraversion–introversion. *J. abnorm. soc. Psychol.* **61**, 52–66.

Brengelmann, J. C. (1960*a*). Extreme response set, drive level and abnormality in questionnaire rigidity. *J. ment. Sci.* **106**, 171–86.

Brengelmann, J. C. (1960*b*). A note on questionnaire rigidity and extreme response set. *J. ment. Sci.* **106**, 187–92.

Brengelmann, J. C. & Brengelmann, L. (1960*a*). Learning and personality. II. FRT reproduction learning analysed by practice stages. *Acta psychol.* **17**, 119–47.

Brengelmann, J. C. & Brengelmann, L. (1960*b*). Learning and personality. III. FRT immediate recall and recognition learning. *Acta psychol.* **17**, 226–52.

Brengelmann, J. C., Hahn, H., Pedley, J. C. & Amato, J. G. (1960). Learning and personality. I. A pilot experiment. *Acta psychol.* **17**, 1113–18.

Brown, R. W. (1953). A determinant of the relationship between rigidity and authoritarianism. *J. abnorm. soc. Psychol.* **48**, 469–76.

Cattell, R. B. (1946). The riddle of perseveration: I. 'Creative effort' and disposition rigidity. II. Solution in terms of personality structure. *J. Personality* **14**, 229–38, 239–67.

Cattell, R. B., Dubin, S. E. & Saunders, D. R. (1954). Verification of hypothesized factors in one hundred and fifteen objective personality designs. *Psychometrika* **19**, 209–30.

Cattell, R. B. & Tiner, L. G. (1949). The varieties of structural rigidity. *J. Personality* **17**, 321–41.

Cowen, E. L. (1952*a*). The influence of varying degrees of psychological stress on problem-solving rigidity. *J. abnorm. soc. Psychol.* **47**, 512–19.

Cowen, E. L. (1952*b*). Stress reduction and problem-solving rigidity. *J. consult. Psychol.* **16**, 425–8.

Cowen, E. L. & Thompson, G. G. (1951). Problem-solving rigidity and personality structure. *J. abnorm. soc. Psychol.* **46**, 165–76.

Cowen, E. L., Wiener, M. & Hess, J. (1953). Generalization of problem-solving rigidity. *J. consult. Psychol.* **17**, 100–3.

Davids, A. (1955). Some personality and intellectual correlates of intolerance of ambiguity. *J. abnorm. soc. Psychol.* **51**, 415–20.

Davidson, P. O., Payne, R. W. & Sloane, R. B. (1964). Introversion, neuroticism and conditioning. *J. abnorm. soc. Psychol.* **68**, 136–43.

Duncker, K. (1945). On problem solving. *Psychol. Monogr.* **58**, no. 5.

Erikson, C. W. & Wechsler, H. (1955). Some effects of experimentally induced anxiety upon discrimination behavior. *J. abnorm. soc. Psychol.* **51**, 458–63.

Eysenck, H. J. (1959). *The Maudsley Personality Inventory*. London: University of London Press.

Eysenck, H. J. (1965). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychol. Bull.* **63**, 258–70.

Field, J. G. & Brengelmann, J. C. (1961). Eyelid conditioning and three personality parameters. *J. abnorm. soc. Psychol.* **63**, 517–23.

Forster, N. C., Vinacke, W. E. & Digman, J. M. (1955). Flexibility and rigidity in a variety of problem situations. *J. abnorm. soc. Psychol.* **50**, 211–16.

French, E. G. (1955). Interrelation among some measures of rigidity under stress and non stress conditions. *J. abnorm. soc. Psychol.* **51**, 114–18.

Goodstein, L. D. (1953). Intellectual rigidity and social attitudes. *J. abnorm. soc. Psychol.* **48**, 345–53.

Gough, H. G. (1952). *The Development of a Rigidity Scale*. Berkeley: University of California Institute of Personality Assessment and Research.

Grant, D. A. & Norris, E. B. (1946). Eyelid conditioning as influenced by the presence of sensitized Beta responses. *J. exp. Psychol.* **37**, 423–33.

Guilford, J. P., Frick, J. W., Christensen, P. R. & Merrifield, P. R. (1957). A factor-analytic study of flexibility in thinking. *Univ. S. Calif. Lab. Report*, no. 18.

Jackson, D. N., Messick, S. J. & Solley, C. M. (1957). How 'rigid' is the 'authoritarian'? *J. abnorm. soc. Psychol.* **54**, 137–40.

Jasper, H. H. (1931). Is perseveration a functional unit participating in all behavior processes? *J. soc. Psychol.* **2**, 28–51.

Kleemeier, R. W. & Dudek, F. J. (1950). A factorial investigation of flexibility. *Educ. psychol. Measur.* **10**, 107–18.

Lankes, W. (1915). Perseveration. *Br. J. Psychol.* **7**, 387–419.

Levitt, E. E. & Zelen, S. (1953). The validity of the *Einstellung* test as a measure of rigidity. *J. abnorm. soc. Psychol.* **48**, 573–80.

Levitt, E. E. & Zuckerman, M. (1959). The Water Jar test revisited; the replication of a review. *Psychol. Rep.* **5**, 365–80.

Luchins, A. S. (1942). Mechanization in problem solving: the effect of *Einstellung*. *Psychol. Monogr.* **54**, no. 6.

Luchins, A. S. (1951). On recent usage of the Einstellung-effect as a test of rigidity, *J. consult. Psychol.* **51**, 81–94,

Luchins, A. S. & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *J. gen. Psychol.* **42**, 279–97.

MAHER, B. A. (1957). Personality, problem solving and the *Einstellung* effect. *J. abnorm. soc. Psychol.* **54**, 70–4.

MALMO, R. B. (1957). Anxiety and behavioral arousal. *Psychol. Rev.* **64**, 276–87.

MALMO, R. B. & DAVIS, J. F. (1961). A monopolar method of measuring palmar conductance. *Am. J. Psychol.* **74**, 106–13.

MOFFITT, J. W. & STAGNER, R. (1956). Perceptual rigidity and closure as functions of anxiety. *J. abnorm. soc. Psychol.* **52**, 354–7.

NIGNIEWITZKY, R. D. (1955). A statistical study of rigidity as a personality variable. Unpublished M.A. thesis, University of London Library.

NIGNIEWITZKY, R. D. (1956). A statistical and experimental investigation of rigidity in relation to personality and social attitudes. Unpublished Ph.D. thesis, University of London Library.

NOTCUTT, B. (1943). Perseveration and fluency. *Br. J. Psychol.* **33**, 200–8.

PALLY, S. (1954). Cognitive rigidity as a function of threat. *J. Personality* **23**, 346–55.

PAYNE, R. W. (1960). Cognitive abnormalities. In *Handbook of Abnormal Psychology*, pp. 193–261. Ed. H. J. Eysenck. London: Pitman.

PERVIN, L. A. (1960). Rigidity in neurosis and general personality functioning. *J. abnorm. soc. Psychol.* **61**, 389–95.

PITCHER, B. & STACEY, C. L. (1954). Is *Einstellung* rigidity a general trait? *J. abnorm. soc. Psychol.* **49**, 3–6.

RAVEN, J. C. (1958). *Guide to Using the Mill Hill Vocabulary Scale with the Progressive Matrices Scales.* London: H. K. Lewis.

RIM, Y. (1953). Perseveration and fluency as measures of introversion–extraversion in abnormal subjects. Unpublished Ph.D. thesis, University of London Library.

ROKEACH, M. (1948). Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.* **43**, 259–78.

ROKEACH, M. (1953). Dogmatism and rigidity as determinants of cognition. Unpublished M.A. thesis, Michigan State University Library.

ROKEACH, M., MCGOVNEY, W. C. & DENNY, M. R. (1955). A distinction between dogmatic and rigid thinking. *J. abnorm. soc. Psychol.* **51**, 87–93.

SCHAIE, K. W. (1955). A test of behavioral rigidity. *J. abnorm. soc. Psychol.* **51**, 604–10.

SCHEIER, I. H. (1954). An evaluation of rigidity factors. *Can. J. Psychol.* **8**, 157–63.

SCHEIER, I. H. & FERGUSON, G. A. (1952). Further factorial studies of tests of rigidity. *Can. J. Psychol.* **6**, 18–30.

SCHMIDT, H. O., FONDA, C. P. & WESLEY, E. L. (1954). A note on consistency of rigidity as a personality variable. *J. consult. Psychol.* **18**, 450.

SHEVACH, B. (1937). Studies in perseveration: VII. Experimental results of tests for sensory perseveration. *J. Psychol.* **3**, 403–27.

SMOCK, C. D. (1955*a*). The influence of psychological stress on the 'intolerance of ambiguity.' *J. abnorm. soc. Psychol.* **50**, 177–82.

SMOCK, C. D. (1955*b*). The influence of stress on the perception of incongruity. *J. abnorm. soc. Psychol.* **50**, 354–6.

SOUEIF, M. I. (1958). Extreme response sets as a measure of intolerance of ambiguity. *Br. J Psychol.* **49**, 329–34.

TAYLOR, J. A. (1953). A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.* **48**, 285–90.

WESLEY, E. L. (1953). Perseverative behavior in a concept-formation task as a function of manifest anxiety and rigidity. *J. abnorm. soc. Psychol.* **48**, 129–34.

WOLPERT, E. A. (1955). A new view of rigidity. *J. abnorm. soc. Psychol.* **51**, 589–94.

ZEAMAN, D., DEANE, G. & WEGNER, N. (1954). Amplitude and latency characteristics of the conditional heart response. *J. Psychol.* **38**, 235–40.

ZELEN, S. L. & LEVITT, E. (1954). Notes on the Wesley rigidity scale. Developments of a short form. *J. abnorm. soc. Psychol.* **49**, 472–3.

(*Manuscript received* 17 *August* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 127–131
Printed in Great Britain

THE INTERACTION OF NOISE AND PERSONALITY WITH CRITICAL FLICKER FUSION PERFORMANCE

By C. D. FRITH

Institute of Psychiatry, University of London

A theory of the interaction of arousal, performance and personality is outlined. On the basis of this theory it is predicted that an increase in noise will improve the performance of extraverts in a c.f.f. task more than that of introverts. An experiment confirming this prediction is described.

In an investigation of the effects of drugs and their interactions with personality it was intended to use critical flicker fusion (c.f.f.) threshold as an indication of the initial level of cortical arousal in a subject and as a measure of the effect of the drug.

As an essential preliminary to this investigation the relation between c.f.f. performance and personality was studied in relation to an arousing agent (auditory noise) whose mode of operation is, theoretically at least, more readily apparent. Arousal is defined in terms of Hebb's (1955) theory. Hebb proposed that afferent impulses serve two quite different functions as a result of the two different pathways to the cortex. One of these is an indirect route through the ascending reticular formation, the non-specific, or diffuse projection system which bombards large areas of the cortex with afferent impulses. This bombardment serves an arousal function, 'toning up' the cortex with a background of supporting activity. Lindsley (1957) has shown that the discrimination of two flashes in the visual cortex of cats can be enhanced by stimulating the reticular formation electrically. This result suggests that c.f.f. thresholds should reflect activity in the reticular formation and thus level of arousal. In line with this hypothesis it is well established that stimulant drugs raise c.f.f. while depressant drugs lower it (Granger, 1960).

Less well established are the relations between c.f.f. and personality. Some studies have shown that extraverts have lower c.f.f. than introverts (Simonson & Brozek, 1952). Eysenck (1963) has put forward a model to explain such individual differences in terms of cortical arousal. He suggests that introverts are characterized by dominance of 'excitatory' processes and extraverts by 'inhibitory' ones, the terms 'excitation' and 'inhibition' being defined by means of clearly demarcated experiments. These concepts are also used in Pavlov's typology and his weak and strong nervous systems are roughly equivalent to introversion and extraversion respectively. Yerkes & Dodson (1908) and many others (e.g. Schlosberg, 1954; Stennett, 1957) have found a curvilinear relation between arousal and performance. A similar curvilinear relation between performance and stimulus intensity (the law of strength and transmarginal inhibition) is an important feature of Pavlov's system and has been interpreted in terms of arousal by Gray (1964). There is also evidence that the optimum level of arousal is lower for 'difficult' than for 'easy' tasks (Broadhurst, 1959).

Introverts, because of the dominance of excitatory processes and their greater sensitivity to incoming stimulation will, under the same conditions, be more aroused than extraverts. If the curvilinear relation between arousal and performance is also

taken into account the relation between personality, arousal and performance will be as shown in Fig. 1. It should be noted that the abscissa in this figure is not the level of arousal, but the extent of arousing factors, i.e. the intensity of external or internal stimulation. This formulation has the advantage that the units on both axes are measured rather than inferred. Such a relation between weak and strong nervous systems has been shown in Teplov's laboratory for reaction time and the critical frequency of flashing phosphene, with stimulus intensity as the arousing factor (Gray, 1964), and also between introverts and extraverts for more complex tasks by Corcoran (1965), arousal being manipulated by incentives and sleep deprivation.

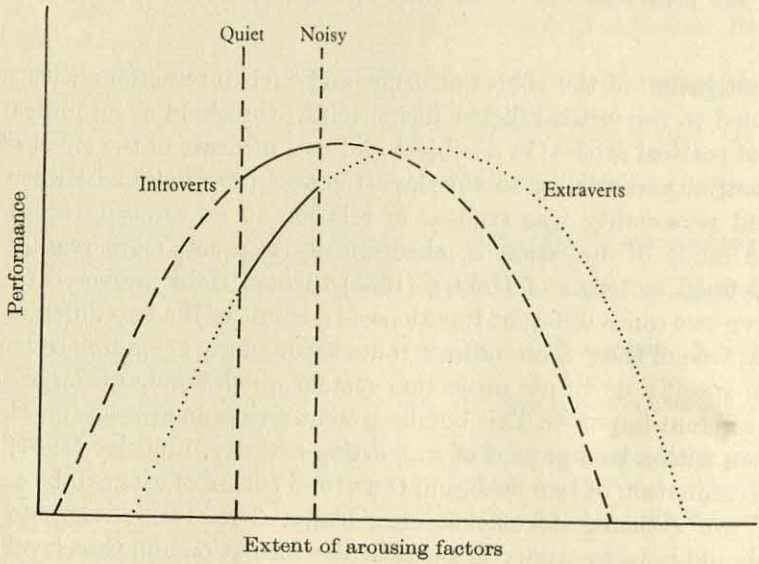


Fig. 1. Hypothesized relationship between the extent of arousing factors, personality and performance.

In line with this model of arousal, personality and performance is a finding of Yermolayeva-Tomina (1964) that with irrelevant distraction (the sound of a metronome) the absolute visual sensitivity of subjects with weak nervous systems decreased while that of those with strong nervous systems increased. These changes tended to reduce the differences between the groups. Claridge (1960) found similar effects with a vigilance task. When a second vigilance task was introduced as distraction, performance on the main task improved in hysterics (neurotic extraverts) and got worse in dysthymics (neurotic introverts).

It seems likely that this model should also apply to c.f.f. performance. Some studies have shown a facilitating effect of auditory noise on c.f.f., others have not (Maier, Bevan & Behar, 1961). These contradictions might be resolved by taking into account initial levels of arousal and personality variables. If c.f.f. performance were to be used as an indication of the effects of drugs it was essential that there should be some prior knowledge of the relationship between arousal, performance and personality for this task. It might be assumed that the task is a relatively easy one and the situation a relatively unarousing one for the subjects. Thus they would be operating below the

optimum level of arousal. Fig. 1 shows that in these conditions an increase in the external arousing factors would produce a greater improvement in the performance of extraverts than in that of introverts. The investigation reported here tested this hypothesis using auditory noise as the arousing factor.

Method

A forced choice technique for measuring c.f.f. threshold was used. Such a technique has been found to be more reliable than the classical psycho-physical methods (Clark, Rutsemann, Link & Brown, 1963) and it also cuts out the effects of differing strategies used by subjects which might well be related to personality. The subjects were required to say which of four lights was flickering. At the threshold area errors rose from zero to 75 %. This threshold area was first found approximately for each subject. Five frequencies, whole numbers of c/s differing by 1 c/s, were then chosen to span the threshold area. Changes in the subject's performance could be measured more sensitively in terms of changes in numbers of errors, rather than in terms of changes in some threshold point measured in c/s. Each testing session lasted 8 min and is described in detail below. Subjects were tested under two noise levels, corresponding to 'quiet' and 'noisy' experimental conditions.

Subjects. There were 40 subjects, 29 male, 11 female. All were normal adult volunteers aged between 18 and 50. The subjects were split into four groups, first, by procedure, half the subjects having the noisy condition and the other half the quiet condition first, and, secondly, by personality, subjects being divided into introverts and extraverts on the basis of their EPI scores (Eysenck & Eysenck, 1964). Details of the four groups are shown in Table 1. Within personality groups the order of presentation of conditions was assigned randomly.

Table 1. *Means and ranges of EPI scores for the four experimental groups*

		Noise–Quiet		Quiet–Noise	
		Mean	Range	Mean	Range
Extraverts	E	15·3	12–18	16·4	12–22
	N	9·9	4–14	7·6	0–21
Introverts	E	7·2	0–11	6·9	3–11
	N	7·4	0–14	9·0	1–16

Apparatus. The flicker was generated electronically. A blocking oscillator provided pulses to an Eccles–Jordan bi-stable multivibrator, which was connected to a Ferranti CL 40 glow modulator tube, which emitted a blue light. The light–dark ratio was 1 : 1 over the range of frequencies used (33–51 c/s). A control panel with four switches produced, for 2 sec, one flickering and three stable lights, each switch producing the flickering light in a different position. The lights were displayed in a square array on a white background, with a central orange fixation point. The white background was a square of side 39 cm, the four lights being centred in it at the corners of a square of side 7 cm. The lights, which were 8 mm in diameter, were numbered anticlockwise, starting at the top left-hand corner. The display was illuminated by a slide projector 100 cm away, illumination of the face and lights being 101·9 ft./Lamberts. The subjects viewed the display at a distance of 160 cm through an artificial pupil 2 mm in diameter, movement of the head being reduced by a chin rest and the necessity to keep all four lights in view. The test field was thus 2 degrees in angular diameter. Noise for the noisy condition was produced by apparatus (power packs, oscillators and electric fans) which gave an average noise level of 50 db. The quiet condition was obtained by the subject wearing a pair of ear guards giving a noise level of 20 db.

Procedure. The approximate threshold was found for each subject under noisy conditions.

Five frequencies differing by 1 c/s were then chosen to fall within the threshold area. In the testing session that followed these frequencies were given to the subject 5 times each in a latin square design. Thus there were five *occasions* on which all frequencies were presented. Within each occasion each frequency was presented five times in quick succession. Thus there were five *presentations* containing all frequencies and all occasions. The exact timing of the session was as follows: five presentations of frequency 1, lights being on for 2 sec and off for 1 sec; 4 sec gap during which frequency was changed; 5 presentations of frequency 2; and so on. Thus the whole session lasted just under 8 min. Subjects were tested thus under both experimental conditions, there being a 5 min rest between sessions. The subject was instructed to call out the number of the flickering light. He was told that if no light appeared to be flickering then he must guess. The position of the flickering light was randomized.

RESULTS

The number of correct responses in each session (125 presentations) was found for each subject. An analysis of variance for these scores is shown in Table 2. From the analysis it can be seen that there was a significant order effect, subjects tending to do better on the second testing session, and a significant interaction between conditions and personality. Table 3 shows that this interaction was due to improved performance in noisy conditions of the extravert group only. The data were also subdivided in terms of conditions, presentations and frequencies but little further information was gained from this. There was a significant difference between frequencies, in that subjects performed better at low frequencies.

Table 2. *Analysis of variance; total number of correct responses in each session*

Source	Sums of squares	D.F.	Mean squares	*F*	*P*
Between subjects	26,076	39	—	—	—
Personality (P)	189	1	189·0	—	—
Groups within P	2,773	2	1,386·5	2·16	—
Subjects within G	23,114	36	641·0	12·44	< 0·001
Within subjects	3,643	40	—	—	—
Order (O)	1,178	1	1,178·0	22·90	< 0·001
Conditions (C)	278	1	278·0	5·40	< 0·05
O × P	63	1	63·0	1·22	—
C × P	270	1	270·0	5·25	< 0·05
Error within	1,854	36	51·5	—	—
Total	29,719	79			

Table 3. *Percentage of correct responses for each personality group under each condition*

	Noisy	Quiet
Extraverts	58·56	52·64
Introverts	53·16	53·12

DISCUSSION

The results confirm the prediction made from the theory relating arousal, performance and personality outlined in the introduction, in that extraverts showed a greater improvement in c.f.f. performance with an increase in auditory noise than did introverts. Table 3 indicates that the introverts showed no change in performance

suggesting that they were already at their optimum level of arousal for this task. Such an interpretation could account for the equivocal results of previous experiments which have investigated individual differences or the effects of auditory noise on c.f.f. Clear differences between personality groups would only appear if the groups were operating well below or above their optimum levels of arousal. The same would apply to the effects of arousing factors. If the arousal of a group of subjects already in the region of the optimum level of arousal were increased then about half would show improved performances and the rest depressed performances. However, the clear increase in performance produced by stimulant drugs is not consistent with this hypothesis. A further complication is the marked order effect found in the present study. This suggests that a learning effect operates to produce changes in performance quite independently of changes in arousal. Quite apart from these complications, if it is correct that in normal conditions subjects are close to their optimum level of arousal for c.f.f. performance, then c.f.f. performance should not be a good measure of changes in arousal, since points from the top of the inverted-U must give the least sensitive and the least reliable reflections of changes in arousal.

The author is indebted to Prof. H. J. Eysenck and Dr G. W. Granger for their help and encouragement, and to the Maudsley and Bethlem Royal research fund for financial assistance.

References

Broadhurst, P. L. (1959). The interaction of task difficulty and motivation: The Yerkes–Dodson Law revived. *Acta Psychol.* **16**, 32–38.

Claridge, G. S. (1960). The excitation-inhibition balance in neurotics. In *Experiments in Personality*, vol. II, ed. H. J. Eysenck. London: Routledge and Kegan Paul.

Clark, W. C., Rutsemann, J., Link, R. & Brown, J. C. (1963). Comparison of flicker fusion thresholds obtained by the methods of forced choice and limits on psychiatric patients. *Percept. motor Skills* **16**, 19–30.

Corcoran, D. W. J. (1965). Personality and the inverted-U relation. *Br. J. Psychol.* **56**, 267–73.

Eysenck, H. J. (1963). *Experiments with Drugs*, chap. 1. London: Pergamon Press.

Eysenck, H. J. & Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London University Press.

Granger, G. W. (1960). Abnormalities of sensory perception. In *Handbook of Abnormal Psychology*, ed. H. J. Eysenck. London: Pitman.

Gray, J. A. (1964). *Pavlov's Typology*, Pt. II. London: Pergamon Press.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychol. Rev.* **62**, 243–53.

Lindsley, D. B. (1957). The reticular system and perceptual discrimination. In *Reticular Formation of the Brain*, ed. H. H. Jasper. London: Churchill.

Maier, B., Bevan, W. & Behar, I. (1961). The effect of auditory stimulation upon the critical flicker frequency for different regions of the visual spectrum. *Am. J. Psychol.* **74**, 67–73.

Schlosberg, H. S. (1954). Three dimensions of emotion. *Psychol. Rev.* **61**, 81–5.

Simonson, R. & Brozek, J. (1952). Flicker fusion frequency: background and applications. *Physiol. Rev.* **32**, 134–78.

Stennett, R. G. (1957). Performance level and level of arousal. *J. exp. Psychol.* **54**, 54–61.

Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *J. comp. neurol. Psychol.* **18**, 458–82.

Yermolayeva-Tomina, L, B. (1964). Concentration of attention and strength of the nervous system. In *Pavlov's Typology*, ed. J. A. Gray. London: Pergamon Press.

(*Manuscript received* 5 *October* 1966;
revised manuscript received 26 *January* 1967)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 133–138
Printed in Great Britain

MUSICAL TASTE AND PERSONALITY

By ELSIE PAYNE
Millfield School, Street, Somerset

An account is given of an experimental investigation into the nature of discriminative musical taste. The aim was to discover the kind of music people enjoy, and in particular, the impact of personality upon taste. The results suggest that personality is an important basic factor in the motivation of taste, though environmental factors strongly modify it, and that a definite if not fully consistent relation exists between classical/romantic tastes and the neuroticism factor of personality. The dimensions are positively related in the sense of classical/romantic taste to stable/neurotic personality.

Research into the psychology of music may be divided into: (*a*) that dealing with *musical ability*, (*b*) that relating to the nature of the *appreciative experience* (*how* one listens to music), and (*c*) that dealing with *musical taste*. The latter may be subdivided into (i) that which is concerned with aesthetic judgement, or good as opposed to bad taste, and (ii) that which explores the type or style of music enjoyed (*what* one likes to listen to). In practice these categories have by no means always been rigidly separated from one another; but recent research, at any rate, has tended to specialize.

(*a*) Seashore (1919, 1938) was the pioneer in the measurement of musical ability; followed later on by Semeonoff (1940) over a wider field. More recently, Wing's Tests of Musical Aptitude (1948, 1962) and Bentley's (1965) tests for younger children have been widely recognized and used. (*b*) Schoen (1923) and Vernon (1930, 1931) did early work on musical appreciation, and Vernon Lee's book, *Music and its Lovers* (1932) did much to stimulate interest in the whole question. Since then work has been done by Burt (1945) and Lawson (1947). Contributions have also been made outside the immediate field of psychology, for example, by Howes (1948, 1958), Susan Langer (1953) and Cooke (1959). (*c*) Vernon (1930) dealt both with the measurement of taste and the question of musical preferences. Semeonoff (1940) in his work on the testing of musical ability also touched on matters of musical judgement. Eysenck (1940, 1941) investigated the nature of aesthetic judgement (his T-factor). (Eysenck's work, like Burt's, did not deal specifically with music but rather with poetry, the visual arts or the arts in general; it is generally agreed, however, that in essentials the arts have much in common with one another.) In extracting his T-factor, Eysenck also extracted two bipolars relating to the type of art enjoyed: (1) the complex/simply constructed, and (2) the highly emotional/more restrained. Lawson (1947) did not touch in detail upon who likes what or why, but she did suggest in this connexion that the tastes of untrained people are completely different from those of musical experts. Cattell & Saunders (1954) reported a substantial enquiry: they set out to find the kind of music people enjoy, and, dealing as Lawson had done with non-musical specialists, based their work on the responses of 196 people to a series of short extracts of music. (The choice of themes, the shortness of each, and the fact that all were played on the piano, not the instrument or instruments for which they were written, are open to criticism from a musician's point of view.) They extracted twelve factors, only eight of which are at all well defined.

It seemed to the writer that of these eight primary factors, several might well come together to form a secondary factor of classicism/romanticism. Eysenck's second factor too might be said to be akin to the classical/romantic dimension; and Vernon (1930) actually used the terms 'classical' and 'romantic' in his investigation. This is particularly relevant to the present enquiry.

The writer has investigated aspects of (*b*) and (*c*) above. The first aspect, dealing with the emotional characteristics of the appreciative experience, has been discussed elsewhere (Payne, 1961). The second aspect, concerning musical taste and personality, is the subject of this paper. Cattell suggested that the chief factors involved in appreciation might have their origin in personality or temperament, and Burt also suggested possible links between personality and various kinds of music. The present enquiry has been confined to the relation between classical/romantic tastes and the neuroticism factor of personality. These results have also been briefly discussed elsewhere (Payne, 1965).

The hypothesis examined was that people with stable/neurotic temperaments tend to have classical/romantic tastes in music. Eysenck's (1960) division of human personality into the three main factors of intelligence, introversion/extroversion and neuroticism (or stability/neuroticism) was adopted. A definition of classicism/romanticism was formulated which was based to a large extent on that of Newton (1962).

The classical/romantic dichotomy is basically a matter of form/feeling. Form may not be lacking in romantic music, but here it is never more than a means to an end—a means to the communication of subjective emotional experience, while in classical music it is an end in itself. Similarly feeling in classical music is the by-product of form, while in romantic music it is the primary instigator of the expression.

The use of the terms 'form' and 'feeling' in connexion with music should perhaps be clarified. 'Form' is used comprehensively to include all aspects of structure, not only large-scale or macrocosmic forms, but also microcosmic patterns, thematic structures for example, and the expansion of these into large-scale forms. 'Feeling' (synonymous here with 'emotion') is used in two senses: to signify, on the one hand, the specific emotional subject-matter of the music (if any) and, on the other, the more general or aesthetic emotional character of the expression which is born of creative enjoyment. It is important to distinguish between these two emotional elements; they are fundamentally distinct from one another, though both may be involved in the same work.

In romantic music the emotional element is more impelling (if not more important) than the overall aesthetic emotion. This emotional subject-matter is not only a specific human emotion, but is concerned specially (to use Newton's terms) with 'mystery, abnormality and conflict'. The emotional element of classical music is necessarily more elusive since it is less apparent. From the point of view of emotion, classical music can be divided into three categories: (1) that in which a small degree of specific emotion penetrates the music indirectly or as a spill-over; (2) that in which a specific emotion is symbolized deliberately and objectively (not subjectively as in romantic music) to produce a 'stylized' expression; and (3) the most 'absolute' of music, that in which specific emotion is non-existent but the aesthetic emotional element is very strong. On this approach the terms 'classical' and 'romantic' can be

applied to music of any period. not merely to the 18th and 19th centuries which were noted for classical and romantic styles respectively.

Method

Tests. The tests used were (*a*) the Maudsley Personality Inventory (Eysenck, 1959) as a test of neuroticism and introversion/extraversion; (*b*) a factual questionnaire, the most important question of which asked testees for lists of (up to six) of their favourite composers (they were also invited to cite specific favourite musical works, which sometimes made classification easier; if, for instance, a person chose Bach's 'St Matthew Passion' as his favourite work, he would be considered more romantic in his tastes than if he had chosen the 'Musical Offering'); and (*c*) Wing's Tests of Musical Aptitude, 1–3.

Table 1. *Classification of composers in terms of the relative 'classicism' (C) and 'romanticism' (R) of their work; the arrows indicate alternative degrees of classicism/romanticism for particular works of some composers, and these scores were used when subjects cited the particular work (see text)*

4*C*,—	3*C*, 1*R*	2*C*, 2*R*	1*C*, 3*R*	—, 4*R*
Corelli		← Beethoven →		Grieg
	Monteverdi →		Dowland	
Vivaldi →		Schubert		Rachmaninoff
	Purcell		Couperin	
Scarlatti →		Debussy		Berlioz
	← Bach →		Dvořak	
Handel →		Ravel		Chopin
	Mozart →		Verdi	
Haydn		Sibelius		Franck
	Mendelssohn →		Tchaikowsky →	
		Vaughan Williams		Schumann
	Brahms		R. Strauss	
		Bartók →		Liszt
	Walton		Elgar →	
		Schönberg →		Puccini
	← Stravinsky →		Berg	
				Mahler
	Tippett →		Shostakovitch	
				Delius
			← Britten	

Before correlations could be calculated between the testees' personalities and their tastes, all the composers quoted had to be assessed in terms of classicism and romanticism. This was done in consultation with a group of expert musicians, and according to the definition of classicism and romanticism which had been adopted. A maximum score of 4 was partitioned into classical and romantic components for each composer: thus Brahms, an intrinsically classical composer writing in a romantic period and hence to some extent romantic in idiom, scored 3*C* plus 1*R*. The main composers were classified as in Table 1; arrows before or after a composer's name indicate that certain of his works are more classical or romantic than the rest. If any of these works were cited specifically, the classical and romantic scores were adjusted accordingly. For example, if Bach's St Matthew Passion were cited—and this alone—the Bach score was 2*C* plus 2*R*.

Subjects. The field of inquiry was restricted to that of *discriminative but non-professional listeners.* (In the researches referred to above, either no distinction was made between discriminative and undiscriminative appreciation, or else, as with Lawson, the focus was definitely upon untrained, and hence probably undiscriminative appreciation.) To be classed as discriminative a testee had *at least* (*a*) to have an extensive listening experience, (both (*a*) and (*b*) assessed by reference to the factual questionnaire), (*b*) to lay claim to be particularly interested in music, and

(*c*) to have an aural score of A, B or C (according to Wing's own classification) in the Wing Tests 1–3 (a modest limit of this sort was considered necessary if, as was assumed, aural ability has some bearing upon musical discrimination). Academic training was not insisted upon, though most of the subjects had had some.

There were 151 subjects altogether. They fell into two main groups: one was of students (first and second year students from a training college, post-graduate music students doing a year's university educational course, a few music specialists from a public school and members of a youth choir and orchestra); the other was of older listeners, from the age of 25 upwards, chiefly from gramophone and music societies. All the students were to some extent trained musicians and all were still studying some aspects of music, but they were not, as yet, professional musicians. The other group included a few who taught music as part of their job, and some who performed, but again there were no full-time professionals and for the most part this was a body of listeners as opposed to performers. The two groups had much in common, but differed in respect of age and degree of musical training.

Assessment of 'taste'. Subjects' tastes were assessed by referring their lists of favourite composers to the above categorization, so as to give each an 'R-score'. If, for example, a testee's list gave $14R$ and $10C$ (for six composers), his R-score would be $(14–10)/6 = +0{\cdot}67$. If, on the other hand, his score was $14C$ and $10R$, his R-score would be $-0{\cdot}67$. Correlations were then calculated between subjects' R-scores and their N (neuroticism) scores.

RESULTS

Correlations between R-scores and N scores were calculated separately for the ST and L groups, as well as for all subjects combined. They were calculated by two methods, by rank order (Spearman's ρ) and by χ^2. Both sets of results are given in Table 2 and the associations between R-scores and N scores are positive in all cases.

Table 2. *Relations between neuroticism (along stable/neurotic) and musical taste (along classicism/romanticism) in a group of students and a group of older listeners to music, and in both groups combined, together with the distributions in broad categories of scores for the two tests*

	Students	Older listeners	All subjects
Sample no.	90	61	151
Neuroticism (M.P.I.):			
Nos. with scores classed as			
Stable	39	32	71
(Neutral)	—	2	2
Neurotic	51	27	78
Classicism/romanticism:			
Nos. whose musical taste (R-score) predominantly			
Classical	27	25	52
(Neutral)	8	12	20
Romantic	55	24	79
Relation between neuroticism (stable/neurotic) and classicism/romanticism			
(*a*) In 2×2 table			
χ^2	14·879	11·665	27·79
P	< 0·001	< 0·01	< 0·001
(*b*) Rank correlation (ρ)	+0·369	+0·290	+0·357
Diff. from zero: t	3·735	2·328	4·665
P	< 0·01	< 0·05	< 0·001

Discussion

There is some difference in significance between the associations as expressed in terms of ρ and χ^2; for the χ^2 tests, however, personality was simply dichotomized into neurotic or stable, and tastes into classical or romantic.

There are probably three reasons why the correlations are not higher. (*a*) In both groups there were a few people whose tastes were in complete opposition to the hypothesis ($N/S = R/S$). Out of the 151 testees there were only seven of these, but sufficient to affect the general pattern of response. In the cases which the writer has been able to investigate, there was some outstanding reason for the discrepancy between personality and taste, and the other discrepancies could doubtless be explained too. (*b*) There seems to be, in any case, a slight pull towards romanticism—regardless of temperament. It is among the young students, between the ages of 17 and 20, that this is most apparent. (*c*) The neuroticism factor of personality is obviously not the only conditioner of tastes.

Factors other than neuroticism which, it was thought, might also influence tastes, have been considered (see Payne, 1965). It was found that sex had no influence at all; age only in so far as there is more of a pull towards romanticism in youth; musical ability, especially aural acuity, very little, except that this presumably extends the field of choice. It has to be remembered, however, that these subjects were, to some extent, homogeneous with regard to aural ability as well as to listening experience and interest in music and indications of the influence of these elements are by no means conclusive. Introversion/extraversion is probably more important. This factor is certainly not itself responsible for classical/romantic tastes. Correlations were calculated between subjects' R-scores and their introversion/extraversion scores, in order to be certain that it was not this factor which corresponded to classical/romantic tastes: these relations were not significant in either group. But introversion/extraversion might still be influential in directing interest towards 'introvert/extravert music' and this might well cut across classical/romantic tastes. A person's academic training seems to have little effect, except again to widen his field of choice. But performing experience is important, as also is the degree and kind of interest that a person has in music.

In spite of the considerable combined power of these influences however, it can be said with some confidence that personality is one of the most powerful and that as far as the factor of neuroticism is concerned, the stable/neurotic dimension of personality is positively related to the classical/romantic dimension in music.

References

Bentley, A. (1965). *Musical Ability in Children and its Measurement*. London: Harrap.

Burt, C. (1945). The psychology of art. In *How the Mind Works*, 2nd ed. London: Allen and Unwin.

Cattell, R. B. & Saunders, D. (1954). Musical preferences and personality diagnosis. *J. soc. Psychol.* **39**, 3–24.

Cooke, D. (1959). *The Language of Music*. London: Oxford University Press.

Eysenck, H. J. (1940). The general factor in aesthetic judgements. *Br. J. Psychol.* **31**, 262–70.

Eysenck, H. J. (1941). Type factors in aesthetic judgement. *Br. J. Psychol.* **31**, 94–102.

Eysenck, H. J. (1959). *Manual of the M.P.I.* London: Oxford University Press.

EYSENCK, H. J. (1960). *The Structure of Human Personality*. London: Macmillan.
HOWES, F. (1948). *Man, Mind and Music*. London: Secker and Warburg.
HOWES, F. (1958). *Music and its Meaning*. London: Athlone Press.
LANGER, S. (1953). *Form and Feeling*. London: Routledge and Kegan Paul.
LAWSON, D. (1947). An analysis of the factors underlying the music-appreciation of untrained adults. Dissertation for Ed. B., University of Glasgow Library.
LEE, V. (1932). *Music and its Lovers*. London: George Allen.
NEWTON, E. (1962). *The Romantic Rebellion*. London: Longmans.
PAYNE, E. (1961). The emotion of music and of music-appreciation. *Music Rev.* **22/1**, 39–50.
PAYNE, E. (1965). Musical taste and personality. *Music Rev.* **26/2**, 129–45.
SCHOEN, M. (1923). Tests of musical feeling and understanding. *J. comp. Psychol.* **5**, 31–52.
SEASHORE, C. E. (1919). *The Psychology of Musical Talent*. New York: Silver, Burdett.
SEASHORE, C. E. (1938). *The Psychology of Music*. New York: McGraw-Hill.
SEMEONOFF, B. (1940*a*). A new approach to the testing of musical ability. *Br. J. Psychol.* **30**, 326–40.
SEMEONOFF, B. (1940*b*). Further developments in a new approach to the testing of musical ability with special reference to groups of secondary school children. *Br. J. Psychol.* **31**, 145–61.
VERNON, P. E. (1930). The phenomena of attention and visualization in the psychology of musical appreciation. *Br. J. Psychol.* **21**, 50–63.
VERNON, P. E. (1930). A method of measuring musical taste. *J. app. Psychol.* **14**, 355–62.
VERNON, P. E. (1931). The psychology of music with special reference to its appreciation, perception and composition. Dissertation for Ph.D. Cambridge.
WING, H. D. (1948). Tests of musical ability and appreciation. *Monograph supplement, Br. J. Psychol.* No. 27.
WING, H. D. (1962). A revision of the Wing Musical Aptitude Tests. *J. Res. mus. Educ.* **10**, no. 1, 39–46.

(*Manuscript received* 5 *February* 1967)

Br. J. Psychol. (1967), **58**, 1 *and* 2, *pp.* 139–142
Printed in Great Britain

INDIVIDUAL DIFFERENCES IN PAIN THRESHOLD AND LEVEL OF AROUSAL

By DIANA R. HASLAM
Department of Psychology, University of Bristol

The pain thresholds of a group of introverts and extraverts were assessed. It was found that the mean pain threshold of the introverts was significantly lower than that of the extraverts. In a further experiment it was found that caffeine, a stimulant drug, lowered the mean pain threshold significantly. The relation between the perception of pain and level of arousal is discussed, and it is argued that the difference in mean pain threshold values as between introverts and extraverts is attributable to a difference in level of arousal.

Hardy, Wolff & Goodell (1943, 1952) suggested that when the pain threshold is measured in 'instructed' subjects there is little variation between individuals in threshold values. (Subjects are termed 'instructed' when a description is given to them of the sensations they are likely to experience.) Other experimenters, however, using the radiant heat technique, have failed to confirm this uniformity (e.g. Slaughter & Wright, 1944; Leduc & Slaughter, 1945).

Since there is evidence of individual differences in pain threshold, it would seem that the relation between pain threshold and personality might be a fruitful area of enquiry. It is one, however, which has received little attention so far, and the results of experiments designed to assess the relation between constitutional–temperamental types and pain threshold in a normal population are inconclusive (Chapman & Jones, 1944; Janoff, Beck & Child, 1950). The present investigation has been concerned, therefore, with the relation of individual differences in pain thresholds to introversion/extraversion rather than to the 'body types and emotional patterns' that have been studied previously.

Experiment I

Method

Subjects

Subjects were selected on the basis of scores obtained on the extraversion scale of the Maudsley Personality Inventory (Eysenck, 1959). Thirty-five students served as subjects, of whom twenty were extraverts (eleven male and nine female) and fifteen were introverts (eight male and seven female).

Apparatus and procedure

A modified Hardy–Wolff (1940) radiant-heat apparatus was used, in which the heat from a 500 w bulb was concentrated by a lens through an aperture 2 cm in diameter. Mains power was supplied to the lamp through a 'Variac' transformer and monitored by means of a voltmeter. The apparatus was calibrated by means of a 'Moll' thermopile and millivoltmeter in terms of mc/sec/cm^2, and the calibration was checked before each experimental session.

For the reasons mentioned by Chapman & Jones (1944, p. 82), a small area of skin in the centre of each subject's forehead was blackened with indian ink. This area was exposed to the radiant heat source for 3 sec at 1 min intervals. (Repeated stimulation of the same area of skin if done in rapid succession will cause lowering of the pain threshold and local hyperalgesia. However, Hardy *et al.* (1952) have shown that with a 1 min interval between successive stimulation of the

same skin area there is no lowering of the pain threshold.) Each subject was told that the procedure was not an endurance test and was not intended to discover how much pain he could stand, but was an experiment to find out at what stage the stimulus was first felt as painful; that is, when he felt a small stab of pain at the end of the 3 sec exposure.

The energy level of the first stimulus was 98 mc/sec/cm², which the majority of subjects described as warm. The intensity of each successive stimulus was increased by approximately 16 mc/sec/cm² since Haslam & Thomas (1967) have found this interval to be optimal for the assessment of pain thresholds, in the sense of giving the most reliable threshold values. The stimulus intensity was increased in this way until the subject reported a small stab of pain at the end of the 3 sec exposure. In assessing pain thresholds it is not possible to carry out the corresponding descending series because of the danger of blistering the skin at high intensities of stimulation.

Before stimulation began, subjects were given a description of the sensations they were likely to experience, and therefore are termed 'instructed subjects'.

Hardy, Goodell & Wolff (1951) reported that skin temperature has a modifying influence on pain threshold in the sense that the higher the skin temperature the lower the pain threshold. Forehead skin temperature was therefore measured at the beginning of the experiment. Room temperature was always between 20·5° C and 22·0° C.

Results

The main results are shown in Table 1.

Table 1. *Mean values of subjects' ages, introversion/extraversion scores, skin temperatures and pain thresholds*

	n	Mean age of subjects in years	Mean introversion/ extraversion score	Mean forehead skin temperature (°C)	Mean forehead pain threshold (mc/sec/cm²)	Range of forehead pain thresholds (mc/sec/cm²)
Introverts	15	24	16	33·8	218	176–264
Extraverts	20	21	33	33·95	249	212–302

Introversion/extraversion

A Mann–Whitney U test (Siegel, 1956) showed that the introverts had a significantly lower forehead pain threshold than the extraverts ($P < 0{\cdot}002$, two-tailed); but differences between the pain thresholds of male and female subjects were not significant for either introverts or extraverts.

Skin temperature

No significant correlation was found between forehead pain threshold and skin temperature for either introverts or extraverts ($\rho = +0{\cdot}47$ and $+0{\cdot}07$ respectively). Unlike the results of Hardy *et al.* (1951), however, the trend was for a high skin temperature to be accompanied by a high pain threshold.

EXPERIMENT II

Experiments have been carried out in the Soviet Union in which assessment was made of the effect of caffeine upon visual and auditory thresholds (Nebylitsyn, 1956, 1961). An experiment is reported here in which the effect of caffeine upon pain threshold was assessed.

Method

Subjects

Twenty-two students served as subjects: fourteen male and eight female.

Apparatus and procedure

The apparatus and procedure for the assessment of pain threshold were the same as those described above.

Each subject took part in one experimental session and one control session. In each session pain threshold was assessed twice, with a 30 min interval between the two assessments, during which subjects sat and read. In the experimental session the assessment of pain threshold was followed by the immediate administration of 6 grains of caffeine citrate; in the control experiment no caffeine citrate was administered. Half the subjects underwent the experimental session first and the other half underwent the control session first. The experimental and control sessions were on different days.

In all sessions care was taken to see that subjects had neither just eaten nor been long without food; also that they had not recently had any alcohol, coffee, tea or drugs. They were not allowed to smoke in the 30 min interval.

Caffeine citrate was used in the present experiment as it is absorbed two to three times more quickly than caffeine. The amount of citric acid present is so small that its effect is thought to be neglibible (personal communication from M. E. Jarvik, 1964).

Results

The effect of caffeine on the mean pain threshold of the total group was to lower it by 14 mc/sec/cm^2 from 223 mc/sec/cm^2 to 209 mc/sec/cm^2. A Wilcoxon matched-pairs signed-ranks test (Siegel, 1956) showed this change to be significant ($P < 0{\cdot}01$). After the administration of caffeine, pain threshold went down in thirteen subjects, up in three subjects and remained unchanged in six subjects. In the control sessions no subject showed any threshold change after the 30 min interval.

Discussion

Arousal is defined by Malmo (1959) as 'the continuum extending from deep sleep at the low activation end to "excited states" at the high activation end'. In defining arousal in this way, Malmo is following Lindsley (1951), whose activation theory of emotion was based largely on findings made with EEG recordings in various states of alertness, and on the discovery of the activating functions of the reticular formation of the brain stem (Moruzzi & Magoun, 1949). The reticular formation not only activates the cortex; it also increases somatic and autonomic activity.

The continuum of arousal suggested by Malmo (1959) might well be extended to include drug-induced anaesthesia at the low activation end, and it seems evident from everyday experience that the relation between arousal and pain perception is of an inverted-U form as follows. Pain is least likely in profound anaesthesia, more likely in superficial anaesthesia and in natural sleep, most likely in a normal waking state, and unlikely in an extremely excited state, as, for instance, in combat and on the playing-field, where it is well-known that wounds and injuries go unnoticed. Hence, pain is likely to be perceived when the level of arousal is neither very low nor very high, but when it is somewhere between these two extremes.

Further evidence that pain perception is related to level of arousal comes from the finding of Expt. II that caffeine lowers the pain threshold significantly. Caffeine gives

rise to an increased level of arousal of the organism as it is a stimulant drug which excites all portions of the central nervous system, in particular the cortex, medulla and spinal cord (Goodman & Gilman, 1955; Sollmann, 1948).

The work of Shagass & Kerenyi (1958) on the sedation threshold indicates that introverts are more aroused than extraverts. It therefore seems possible to explain the finding that introverts have a lower pain threshold than extraverts on the ground that they function at characteristically different levels of arousal.

REFERENCES

CHAPMAN, W. P. & JONES, C. M. (1944). Variations in cutaneous and visceral pain sensitivity in normal subjects. *J. clin. Invest.* **23**, 81–91.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. University of London Press.

GOODMAN, L. S. & GILMAN, A. (1955). *The Pharmacological Basis of Therapeutics*. New York: Macmillan.

HARDY, J. D., GOODELL, H. & WOLFF, H. G. (1951). The influence of skin temperature upon the pain threshold as evoked by thermal radiation. *Science, N.Y.* **114**, 149–50.

HARDY, J. D., WOLFF, H. G. & GOODELL, H. (1940). Studies on pain: A new method for measuring pain threshold: observations of spatial summation of pain. *J. clin. Invest.* **19**, 649–57.

HARDY, J. D., WOLFF, H. G. & GOODELL, H. (1943). The pain threshold in man. *Proc. Ass. Res. nerv. Dis.* **23**, 1–15.

HARDY, J. D., WOLFF, H. G. & GOODELL, H. (1952). *Pain Sensations and Reactions*. Baltimore: Williams and Wilkins.

HASLAM, D. R. & THOMAS, E. A. C. (1967). An optimum interval in the assessment of pain threshold. *Q. J. exp. Psychol.* **19**, 54–8.

JANOFF, I. Z., BECK, L. H. & CHILD, I. L. (1950). The relation of somatotype to reaction time, resistance to pain and expressive movement. *J. Personality* **18**, 454–60.

LEDUC, E. H. & SLAUGHTER, D. (1945). The effects of a combination of acetanilid, caffeine and sodium bromide on the pain threshold in man. *Anesth. Analg.* **24**, 147–54.

LINDSLEY, D. B. (1951). Emotion. In *Handbook of Experimental Psychology*, ed. S. S. Stevens. New York: Wiley.

MALMO, R. B. (1959). Activation: a neuropsychological dimension. *Psychol. Rev.* **66**, 367–86.

MORUZZI, G. & MAGOUN, H. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroenceph. clin. Neurophysiol.* **1**, 455–73.

NEBYLITSYN, V. D. (1956). The relationship between sensitivity and strength of the nervous system. In *Typological Features of Higher Nervous Activity of Man*. Moscow.

NEBYLITSYN, V. D. (1961). Individual differences in the strength and sensitivity of both visual and auditory analysers. In *Recent Soviet Psychology*, ed. N. O'Connor. Oxford: Pergamon.

SHAGASS, C. & KERENYI, A. B. (1958). Neurophysiologic studies of personality. *J. nerv. ment. Dis.* **126**, 141–7.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

SLAUGHTER, D. & WRIGHT, T. (1944). A modification of the Hardy–Wolff–Goodell pain-threshold apparatus. *Anesth. Analg.* **23**, 115–19.

SOLLMANN, T. (1948). *A Manual of Pharmacology and Its Applications to Therapeutics and Toxicology*. Philadelphia: Saunders.

(*Manuscript received* 26 *January* 1967)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 143–145
Printed in Great Britain

NOTE ON THE DISTRIBUTION OF DOMINANCE TIMES IN BINOCULAR RIVALRY

By W. J. M. LEVELT*

Institute for Perception RVO–TNO, Soesterberg, The Netherlands

The dominance periods of a stimulus in binocular rivalry show a characteristic time distribution that gives important clues as to the underlying mechanism in alternation. It is shown that the distribution can be approximated by a Γ-function, which turns out to have a positive integral exponent. With an integral exponent the function describes a Poisson distribution. This suggests the existence of an underlying dominance generating process that is *discrete* in nature. The parameters of this process are determined by properties of the recessive stimulus in the other eye. The discrete events may be 'flicks' of eye movement.

This note is an appendix to the writer's paper in the previous issue of this *Journal* (Levelt, 1966). The reader is referred to that paper for the experiments on which the present discussion is based and also for more precise definitions of the symbols used. The main conclusion of that article was that in binocular rivalry the mean duration of the predominance period of one eye is a function only of the stimulus strength in the other eye.

Further insight in how the shift in eye dominance is generated may be obtained by considering the distribution of dominance times (t_l) of, say, the left eye during rivalry. This analysis is based on the alternation recordings of ten subjects for two different stimulus conditions. As a first step, means ($\bar{t}_l$) and standard deviations (s_l) were calculated over the sample of dominance times of each of the ten subjects under the two stimulus conditions. Relations between means and standard deviations were shown in Figs. 6 and 7 of Levelt (1966) and they can be summarized in both cases by the expression $\bar{t}_l = 2s_l$. This relation holds fairly well for both large and small t_l. It at once excludes the possibility that dominance times can be described by an exponential distribution of 'holding times', for, if the chance of no left–right shift during t is $P_o(t) = e^{-\lambda t}$, then the mean time $\bar{t} = 1/\lambda$, and also $\sigma = 1/\lambda$. This does not agree with the findings.

Because of the constant ratio, $\bar{t}_l/s_l$, the t-values for each subject in each of the two experiments were divided by the mean ($\bar{t}$) of these values in order to put all twenty distributions on the same scale. Hence all of the rescaled distributions have the same mean $\bar{t} = 1$. From these values a single histogram of the t_l-distribution was made for both experiments and all subjects (see Fig. 1).

The distribution of the t_l-values can be approximated by a function of the form $\phi(t) = t^{\alpha-1}\, e^{-t}/\Gamma\alpha$, in which

$$\Gamma\alpha = \int_0^\infty t^{\alpha-1}\, e^{-t} dt$$

(for positive integers $\Gamma\alpha = (\alpha-1)!$). The t-distribution, $f(t)$, for each subject, and also the distribution of Fig. 1, may be conceived of as deformations of $\phi(t)$, with

* Now at the Department of Psychology, University of Groningen, The Netherlands.

$f(t) = \lambda\phi(\lambda t)$ in which λ represents a scale factor. The mean of the t-distribution is

$$\mu = \int_0^\infty \{t^\alpha e^{-t}/\Gamma\alpha\}\, dt = \alpha,$$

the variance is

$$\sigma^2 = \int_0^\infty \{t^{\alpha+1} e^{-t}/\Gamma\alpha\}\, dt - \mu^2 = \alpha,$$

hence the standard deviation $\sigma = \sqrt{\alpha}$.

The experimental requirement that $\mu = 2\sigma$ is met by $\alpha = 4$ ($\mu = 4$, $\sigma = 2$). Hence the best fitting $\phi(t) = t^3 e^{-t}/3\,!$, and

$$f(t) = \lambda(\lambda t)^3 e^{-\lambda t}/3\,!. \tag{1}$$

The mean of this distribution is $\bar{t} = 4/\lambda$, whereas its standard deviation is $2/\lambda$. This function has been fitted to the histogram in Fig. 1. The area under the curve was

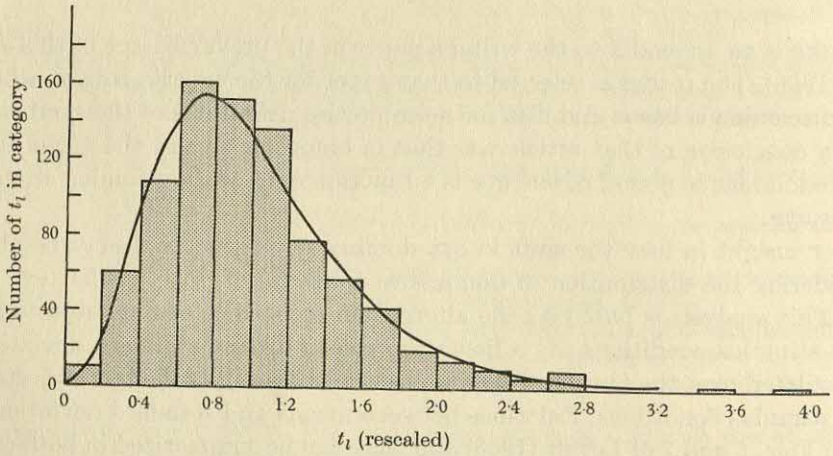


Fig. 1. Histogram of t_l-distribution ($\bar{t}_l = 1$), with best fitting density function $\phi(t) = t^3 e^{-t}/6$.

made equal to the area under the histogram, and a scale factor $\lambda = 4$ had to be used in order to give $\bar{t} = 1$, since the histogram had been rescaled to have a mean of unity. The approximation is fair, but it is realized that other functions may fit as well. The question of interest, however, is whether the function can be understood as an expression of some underlying mechanism.

In the previous paper (Levelt, 1966) it was shown that $\bar{t}_l$ is independent of what was there called λ_l, the strength of the left eye stimulus. It is not accidental that the same term λ has been used above. The scale factor λ in expression (1) may be taken to be the strength of the stimulus presented to the other eye, and evidence was indeed presented that $\bar{t}_l$ and the *right* eye stimulus strength, λ_r, were related by a monotonic decreasing function $\bar{t}_l = f(\lambda_r)$ (see Levelt, 1966, p. 227). It is now possible to specify this monotonic decreasing function as $\bar{t}_l = 4/\lambda_r$, in which an increase of λ_r implies a reduction of $\bar{t}_l$.

One possible mechanism that could account for an effect of stimulus strength so that a function of form (1) results, is a summative one in which the recessive stimulus produces a series of randomly distributed 'excitation spikes'. If the chance of no

spike occurring in t is $P_o(t) = e^{-\lambda t}$, the waiting time for one spike has a probability density function $f_1(t) = \lambda e^{-\lambda t}$. The waiting time up to the nth spike can be found by applying

$$f_n(t) = \int_{-\infty}^{\infty} f_{n-1}(u)\, f_1(t-u)\, du$$

(on the assumption that the waiting time for the next spike is independent of the previous waiting times). For $n = 4$, $f_4(t) = \lambda(\lambda t)^3 e^{-\lambda t}/3!$ which is identical with expression (1), above.

This may be interpreted in the sense that the summative effect of four successive spikes from the recessive stimulus is necessary and sufficient to re-establish dominance for that stimulus. The symbol λ now represents simply the number of spikes per unit time. The experiments gave values of $\bar{t}_l$ of about 2·6 sec. Taking $\bar{t}_l = 4/\lambda_r = 2{\cdot}6$, gives $\lambda = 1{\cdot}54$ sec or an average interspike interval of about 0·65 sec. However, there are large individual differences in λ_r. This interspike time might be associated with 'flicks' in eye movement, suggesting some summating effect of time contours. It is not possible at present to give a more psychophysical account of such a summative process, and further elaboration of the term 'excitation spike' as used in this note would therefore be premature.

References

Levelt, W. J. M. (1966). The alternation process in binocular rivalry. *Br. J. Psychol.* **57**, 225–38.

(*Manuscript received* 19 *May* 1967)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 147–151
Printed in Great Britain

A NOTE ON INTELLIGENCE AND SEASON OF BIRTH

By GÖSTA W. BERGLUND
Institute of Education, University of Uppsala, Sweden

The purpose of this study was to investigate whether the relation between intelligence and season of birth, reported by some British and other workers, can be confirmed (1) in a sample of 708 normal Swedish children and (2) in a sample of 237 Swedish children below normal intelligence. In the normal sample, no significant difference in intelligence was obtained between groups of children born during different parts of the year. In the sample of backward children a trend was observed; the mean I.Q. was highest in the group consisting of children born during the autumn.

Differential psychology, as is well known, contains many problems. Some of them may be regarded as great and vital, others as small and unimportant. The problem of the relation between season of birth and intelligence is usually included in the latter category. Nevertheless, this problem has, during recent years, become increasingly topical, and has aroused much interest in its psychological and pedagogical aspects, especially among some British workers.

By far the most comprehensive investigation in this field was made by Pintner & Forlano (1933). They summarized their data as follows: 'A study of the data by months, seasons and by various groups of months, shows consistently a lowest mean I.Q. for the winter months (January to March). The mean difference in I.Q. between winter and the highest seasonal mean is statistically reliable' (Pintner & Forlano, 1933, p. 583). They advanced the view that this effect is due to influence of climate during the first months after birth. Similar results were arrived at by Fitt (1941), who came to the conclusion that persons born during the summer/autumn half year tend to be superior both physically and mentally. Fitt did not ascribe the observed effect to the season of or immediately after birth, but rather to the season when conception took place.

During recent years, a British worker, Orme (1962, 1963), has devoted much interest to the problem. The latter study was based on a sample of 188 backward adults within an I.Q. range of 40–69. Persons born during summer and autumn had significantly higher I.Q.s than persons born in winter and spring. Persons born in summer and winter had the highest and lowest I.Q.'s respectively. Orme does not consider that this variation is directly related to the difference in temperature between the two seasons. The result suggests rather that the effect is related to climatic changes during the whole prenatal period. Persons born in summer are not superior in intelligence owing to high summer temperatures, but because of steadily rising temperature during the prenatal period, through spring to summer. This hypothesis was later confirmed in a group of supernormal subjects (Orme, 1965). In order to check whether Orme's results could be confirmed on a group of normally gifted adults, Davies (1964) studied a sample of 300 men aged 20–70 years. Two types of intelligence test were administered. No significant differences were found between the summer/autumn group and the winter/spring group. Not even when summer was combined with

spring and winter with autumn could any statistically reliable differences be observed. (See also Crookes, 1963.)

The purpose of an analysis of the samples presented below is to study whether the differences observed by Orme and others can be found (1) in a group of normal children, and (2) in a group of backward children. In conjunction with one of his investigations, Orme (1963, pp. 275–276) stated that 'similar studies to the present one in countries with different mean levels of temperature and with different seasonal variations of temperature would be of great interest. . . .The present findings suggest one area worthy of more attention.'

PROBLEMS

(1) *Is there, in normally intelligent children, any relation between intelligence and season of birth?*

Subjects and tests

The experimental sample comprised 708 children born in 1954. When the investigation was made, all the subjects were pupils in grade 4 (normal class) at schools in Uppsala, and at some schools in the surrounding rural areas. A survey of all subjects according to sex and birth month is given in Table 1. All the subjects were tested by the Terman–Merrill scale, Form L. Half of them were tested during the autumn term of 1964, and the rest during the following spring of 1965.

Table 1. *Number of subjects in the normal sample*

	Birth month												
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
Boys	34	27	40	33	44	22	25	33	29	28	35	32	382
Girls	39	18	32	21	30	21	26	24	25	30	34	26	326
Total	73	45	72	54	74	43	51	57	54	58	69	58	708

Table 2. *Mean monthly temperatures (° C) in Uppsala and the surrounding district*

	Month											
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1953	−3·6	−4·6	+2·6	+6·0	+10·0	+17·0	+16·4	+15·3	+10·9	+8·8	+2·9	+1·5
1954	−5·4	−7·7	+0·1	+2·6	+11·1	+14·2	+15·7	+14·9	+11·0	+6·8	+0·8	+1·1

Results

The mean monthly temperatures for the years 1953 and 1954 at Uppsala and in the surrounding district are given in Table 2. December, January and February had the lowest temperatures and June, July and August had the highest temperatures. With reference to this table the sample was divided into four groups. Subjects born in December, January and February (winter) were combined to make one group. The second group consisted of pupils born in March, April, May (spring), the third group of pupils born in June, July, August (summer) and the fourth group of those born in September, October, November (autumn).

The I.Q.'s of all the subjects were determined. Different measures were then calculated for a description of the I.Q. distribution for each of the four groups and for the

sample as a whole. A summary of the data is presented in Table 3 which shows, for each season of birth and for the sample as a whole: median I.Q. (Md), upper and lower quartiles (Q_3 and Q_1), quartile deviation (Q.D.), mean I.Q. (M), standard deviation (S.D.) and number of subjects (n).

Table 3. *Measures for a description of the I.Q. distribution for each season of birth and for the sample as a whole*

	Winter	Spring	Summer	Autumn	Whole sample
n	176	200	151	181	708
Md	110·5	109·7	112·1	109·7	110·6
Q_1	100·5	100·6	101·9	99·8	100·5
Q_3	121·5	123·8	123·5	124·6	123·2
Q.D.	10·5	11·6	10·8	12·4	11·6
Mean	111·6	112·6	113·0	112·2	112·3
S.D.	16·5	17·9	16·4	18·1	17·5

First a distribution-free method (Peatman, 1964, p. 376) was used for the analysis of the significance of differences between the groups; the value of χ^2 was not statistically significant ($\chi^2 = 1{\cdot}84$; D.F. 3; $P > 0{\cdot}5$). Since the skewness of the distributions may be regarded as moderate, a parametric method, the t test, was then applied to test the null hypothesis of no differences between the means of the four groups. None of the t values was significant at $P \leqslant 0{\cdot}05$. For the difference in I.Q. between the highest and the lowest seasonal mean (summer and winter respectively) $t = 0{\cdot}78$ ($P > 0{\cdot}10$). Thus no statistically reliable differences exist in respect of intelligence for subjects born during different parts of the year, and Orme's observations could not be verified in this sample. It is also to be observed that the results presented above are based on a sample of normally intelligent subjects. The same is true of the studies of Davies and Crookes, reported above. Orme's samples, on the other hand comprised groups of subnormal and supernormal ability. Is it possible that the relation obtained by Orme can be observed only at the extremes of the normal curve?

(2) *Is there, among backward children, any relation between intelligence and season of birth?*

Subjects and tests

This experimental sample comprised 237 children, pupils of schools at Uppsala. At the time of testing, all the subjects were about to be transferred to special classes. The number of subjects divided according to birth month, is shown in Table 4. The Terman–Merrill scale, Form L was administered to all the subjects. All tests were administered by the school psychologist.

Results

In order to obtain a reasonably large number of subjects in each group, subjects born in December, January, February and March were combined to make one group. The second group consisted of pupils born in April, May, June, and July and the third group of those born in August, September, October and November. A summary of the measures calculated for each of the three I.Q. distributions and for the total sample is given in Table 5. The same tests of significance were used as in the analysis of the normal sample.

A clear trend can be observed in this case. The median (Md) and the mean (M) for the autumn group are higher than the corresponding values for the winter and summer groups ($\chi^2 = 4{\cdot}64$; D.F. 2; $0{\cdot}10 > P > 0{\cdot}05$). The t test was also used to determine the reliability of the obtained differences. Table 6 summarizes the relevant data and demonstrates that the difference in mean I.Q. between autumn and winter is statistically significant at the 5% level.

Table 4. *Number of subjects in the sample of backward children*

Birth month												
Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
11	17	27	21	16	11	20	23	22	25	26	18	237

Table 5. *Measures describing the I.Q. distribution for each of the three groups of backward children and the sample as a whole*

	Dec.–Mar.	Apr.–July	Aug.–Nov.	Whole sample
n	73	68	96	237
Md	77·6	78·5	79·8	78·7
Q_1	72·5	73·4	76·2	74·4
Q_3	82·1	83·5	83·8	83·3
Q.D.	4·8	5·1	3·8	4·5
Mean	77·4	78·4	79·9	78·7
S.D.	7·7	8·2	6·8	7·5

Table 6. *t-values for significance of differences between group means*

Groups compared	t	
Dec.–Mar. *v.* Apr.–July	0·74	$P > 0{\cdot}10$
Dec.–Mar. *v.* Aug.–Nov.	2·19	$0{\cdot}05 > P > 0{\cdot}01$
Apr.–July *v.* Aug.–Nov.	1·23	$P > 0{\cdot}10$

DISCUSSION

The explanation of the results obtained in the low-intelligence group may be a simple one. There are many reasons why a pupil may be assigned to a special class. The transfer to special classes is mainly made on the basis of school achievement and intelligence test scores together with the results from a short medical examination. In most cases the special class placement primarily becomes a placement on demonstrated achievement. The strongest reason for children's failure to achieve may be an unsatisfactory mental age, and the same mental age results, as is well known, in lower I.Q.'s for children born early in the year than for children born during the autumn, who are the youngest pupils in a year group. Thus the results obtained in this sample may be the consequence of a restricted range (a greater homogeneity) in mental age.

The findings of studies made from the pedagogical aspect are to a certain extent in contrast with the results of studies concerned with the relation of intelligence to season of birth. John (1962) showed that among 1300 children, retarded in reading, there was a significantly higher proportion of children born in summer. Williams (1959) arrived

at the same result, and also showed that in a group of 1511 children assigned to special classes, there was a clear majority of children born during summer. This was confirmed in a later report, and Williams (1964, p. 253) says: 'It seems fairly clear that for the 265 children who have been attending special schools in South Wales a disproportionate number have Summer birthdays.' The problem may be a larger and more intricate one than has hitherto been considered in psychological and pedagogical circles. In the writer's opinion, the explanation of differences in intelligence between groups of children born during different parts of the year in terms of climatic factors must be subject to doubt. Further research is needed to arrive at a complete solution of the problem.

References

Crookes, T. G. (1963). A note on intelligence and date of birth. *Br. J. med. Psychol.* **36**, 355–6.

Davies, A. D. M. (1964). Season of birth, intelligence and personality measures. *Br. J. Psychol.* **55**, 475–6.

Fitt, A. B. (1941). *Seasonal Influences on Growth, Function and Inheritance.* Oxford: Oxford University Press.

John, E. (1962). *The Age Factor in Reading Retardation Researches and Studies*, no. 24, pp. 1–7. University of Leeds, Institute of Education.

Orme, J. E. (1962). Intelligence and season of birth. *Br. J. med. Psychol.* **35**, 233–4.

Orme, J. E. (1963). Intelligence, season of birth and climatic temperature. *Br. J. Psychol.* **54**, 273–6.

Orme, J. E. (1965). Ability and season of birth. *Br. J. Psychol.* **56**, 471–5.

Peatman, J. G. (1964). *Introduction to Applied Statistics.* New York: Harper and Row.

Pintner, R. & Forlano, G. (1933). The influence of month of birth on intelligence quotients. *J. educ. Psychol.* **24**, 561–84.

Williams, P. (1959). The Growth of Reading Vocabulary. Unpublished Ph.D. Thesis, University of London, University Library.

Williams, P. (1964). Date of birth, backwardness and educational organization. *Br. J. educ. Psychol.* **34**, 247–55.

(*Manuscript received* 14 *November* 1966)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 153–162
Printed in Great Britain

THE GENETIC DETERMINATION OF INTELLIGENCE: A REPLY

By CYRIL BURT
University College London

I. The assessment of intelligence

I am most grateful to Prof. Stott and Dr Lewis, not only for their appreciative comments on the studies of monozygotic twins carried out by my coworkers and myself, but also for the lucid way in which they have summarized their objections to the conclusions that we drew. The criticisms which they now advance have in point of fact already been put forward in earlier papers which they published elsewhere (Lewis, 1957; Stott, 1959); and I endeavoured at the time to meet their various objections (cf. Burt, 1957, 1960; Stott appended a brief comment to this last article, but ignored my evidence on the prenatal conditions of twins). In their latest papers (Stott, 1966; Lewis, 1966) they neither mention my counter-criticisms nor attempt to rebut the arguments and the further evidence which I brought forward in my replies.

Dr Stott concedes that the methods which my colleagues and I employed have enabled us to 'overcome many of the former objections' to research on twins. The fact that 'there is some genetic reason why the intelligence scores of twins tend to resemble one another', he says, now seems 'inescapable'. What he questions, therefore, is 'Burt's conclusion that "individual differences in intelligence are influenced *far more* by genetic constitution than by post-natal or environmental conditions"', i.e. not the fact of inheritance, but the importance attached to it. In weighing his arguments and ours, however, much will depend on the meaning assigned to the term 'intelligence'. Unfortunately, in quoting my conclusion Stott has omitted the explanatory phrase which I inserted. I put the word 'intelligence' between inverted commas, and added 'particularly when the assessments have been carefully checked'. By this I intended to remind the reader that, as indicated at the outset of my paper, what I was discussing was not 'intelligence' in the popular sense (which usually includes acquired knowledge and skill, and an appropriate 'conceptual equipment', to use Stott's own interpretation), but rather the psychologist's attempts to assess the individual's 'innate general ability'—a purely 'hypothetical factor'. My object was to demonstrate that, when these assessments were reached, not by taking scores on some familiar 'group test' just as they stand, but by adopting the more elaborate procedures which my colleagues and I had used, then the errors of assessment resulting from differences in environmental opportunity were comparatively slight. This conclusion, if sustained, has an obvious practical importance. It would go far to dispose of doubts and difficulties like those expressed by the writers of the Plowden report in their discussion both on the value of the I.Q. as currently assessed and of the possibility of identifying gifted children at an early stage in their school career (Plowden, 1967, pp. 20, 305).

Stott begins by arguing that the interpretation of the results reported in my paper must vary with the 'conceptual framework' that is adopted. The 'conceptual framework' which I myself assumed was that of most contemporary geneticists; and this, it seemed to me, presented a strong antecedent presumption in favour of the view that differences in genetic constitution would produce differences in mental capacity as wide as those in bodily height, and that these mental differences would be determined by much the same types of unifactorial and multifactorial inheritance as are known to determine physical differences (cf. Burt, 1912; Burt & Howard, 1956, and refs.). Stott's 'conceptual framework' appears to be that of the contemporary environmentalist. Of course every geneticist recognizes that environment exerts *some* influence, and that, so far as *ultimate* achievements are concerned, it may be of crucial importance. The thoroughgoing egalitarian, however, wants to go further than this. He maintains, as Watson did, that practically *all* observable differences in mental performance are the effects of learning and of motivation operating under varying environmental conditions. Stott does not entirely deny the possibility of innate differences in mental capacity; but (if I understand him rightly) he holds that, in a society like our own, they are of little practical or educational importance.

II. ALTERNATIVE EXPLANATIONS

His general criticism of my conclusions turns on the fact that they 'do not emerge as *necessary* conclusions from [my] findings'. There are, he contends, alternative interpretations, which, as an environmentalist, he would prefer to adopt. He repeatedly complains that I 'make no reference' in my paper to this or that hypothetical explanation, which would render a genetic explanation redundant. Such complaints would no doubt be serious were the paper he criticizes the only one I had published on the subject. But, if he will look again at the various publications referred to at the end of my article, he will find each of these alternatives examined at considerable length.

The greater part of his paper is concerned with criticisms of the three 'major assumptions' which he believes I have explicitly or tacitly adopted, for each of which he wishes to substitute an alternative assumption of his own.

1. His first objection is levelled against what he calls my 'profession of faith in an "innate, general, cognitive factor"'. He supposes that I began by observing 'individual differences in test scores', and then, in order to account for them, 'proposed that there is some distinctive ability termed intelligence'. He thus presumes that I followed the popular idea of the methodology appropriate to an inductive science—observation first, hypothetical generalization afterwards. But that, as I have repeatedly insisted (e.g. Burt, 1958*a*), inverts the true procedure. The alternative hypotheses should be formulated first, and the observational studies should then be planned in the hope of confirming one of the hypotheses, and disproving the rest.

This procedure, as will be obvious from my paper on 'The concept of intelligence' (Burt, 1955), was the one that I followed in my various studies of inheritance; and I expressly asked that 'the logic of the argument should be carefully noted: as in all natural science, the mode of proof is indirect; hence the conclusions can never be "necessary" or "certain", but only probable. The critic', I added, 'commonly misses

this point, and revels in demonstrating that some alternative can be conceived. But one can always think up alternatives; a hypothesis empirically shown to be probable can only be overthrown by proving that the proposed alternative *is still more probable*'. This Stott failed to do. Most of his contentions are mere armchair arguments suggested by analogies with other inquiries, but never supported by any *ad hoc* investigation directly designed to refute my own conclusions.

The hypothesis which I set out to test was threefold. It postulated a theoretical component or 'factor' which was (i) cognitive as distinct from conative, (ii) general and so distinct from the narrower 'group factors' underlying special abilities, and (iii) innate rather than acquired. The alternative hypotheses denied one or more of these assumptions. The earlier critics repudiated the notion of a general factor, and recognized the existence of group factors (or 'primary abilities') only. This, however, is an alternative that neither Stott nor Lewis wishes to resurrect. The assumption which forms the target of Stott's opening attack is the notion that the factor is 'cognitive', or, as I put it in discussing Lewis's criticisms, 'cybernetic' rather than 'dynamic'. 'It is not necessary', he writes, 'to attribute individual variation to differences in some cognitive factor . . . There is nothing in Burt's results to contradict the view that the genetic factor making twins more similar in "intelligence" is a motivational one'.

I admit that the attribution is not 'necessary', merely 'more probable'; and I readily concede that the results for twins do not of themselves demonstrate that the 'genetic factor' is 'cognitive rather than conative': that was not an issue with which the article was concerned. It had already been dealt with in previous publications. In a series of experiments on motivation, Williams, Flugel, and I offered monetary rewards to children who could beat their own records in a wide variety of mental and scholastic tests. We found that this extra incentive had a marked influence on speed tests (e.g. simple addition with a time-limit), a smaller but still an appreciable influence on success in the kind of tests used in junior scholarship examinations and the 11 plus, including most group tests of intelligence (not the reasoning tests, however), and no discernible influence at all on individual tests of the Terman–Binet type (Burt & Williams, 1962).

Stott, however, has in mind, not only the child's motivation during the actual tests, but also 'his intrinsic motivation during the early stages of development'—his 'curiosity', 'eagerness', and similar character-qualities, which (if I understand him rightly) he regards as the effect partly of 'habit' and partly of 'operative genetic factors'. The suggestion is by no means new. In an earlier article (Burt, 1955, pp. 165 f., section on 'The Factor as Cognitive') I noted that 'a number of writers argue that the factor common to mental activities tested is not cognitive but conative —an interpretation which has a warm appeal for those who cherish the doctrine of intellectual equality'; and I explained how Maxwell Garnett (1911) had interpreted Spearman's theory of 'mental energy' as meaning that the factor was essentially motivational. In reply I cited the evidence provided by correlation tables which included motivational traits as well as cognitive (e.g. Burt, 1946). These plainly indicated *two* broad factors—one covering 'intellectual and practical capacities', the other 'affective and conative characteristics', with little or no overlap.

Stott ignores this evidence, and roundly declares that 'means are lacking to

distinguish motivations from hypothetical cognitive ability; there is no convenient phenotype to present to the geneticists, such as colour of eyes or hair: all the data are behavioural'. Actually the initial 'datum' on which my hypothesis was based was not 'behavioural', but physiological, namely, the familiar distinction between the biochemical energy needed for both mental and muscular activity and the physical structure of the brain as a device for controlling and directing that energy into the appropriate channels—a distinction drawn by Sherrington in his account of the 'integrative action' of the nervous system. Bolton's microscopical studies of the brains of normal and feeble-minded individuals revealed clearly observable differences in their anatomical structure. In the brains of defectives the nerve-cells were relatively few in number, had fewer branches, and were irregularly arranged; and these peculiarities, so he found, were, as a rule, characteristic of all the areas and each of the layers of the cerebral cortex. This was my 'phenotype'. And on this basis I proceeded to argue that these observable differences in anatomical structure, like other anatomical differences, would be in all probability genetically determined (Burt, 1911, pp. 95 f.). The correlational evidence was collected specifically to confirm this inference.

2. My 'second major assumption', according to Stott, is that 'the contributions of heredity and environment are relatively constant irrespective of the population studied'. To disprove it he quotes Gordon's investigation of canal-boat children in which 'the children appeared near-feebleminded, judging by their scores on test-tasks culturally foreign to them, but normal or nearly so, when tested with performance tests'. I can find nothing in my article or elsewhere suggestive of the 'major assumption' thus ascribed to me. In discussing 'the theory of multifactorial inheritance and its application to intelligence' Miss Howard and I expressly emphasized that our conclusions about the relative contributions of heredity and environment 'would *not* hold good of a population of a different genetic composition or at a different cultural level'. Gordon's research, among several others, was actually quoted as illustrating the need for 'care in the choice of tests and in checking the results', particularly when dealing with exceptional groups (Burt & Howard, 1957, pp. 61 f.). Why Stott should suppose that my own 'test-tasks' were also 'culturally foreign' to the children tested I cannot imagine.

3. My 'third major premiss', he says, is that 'inferences about twin development can be extended to non-twins' by a mere theoretical 'extrapolation'. This again reverses the order of our investigations. My colleagues and I began (Burt, 1912) with correlational studies of non-twins (brothers and sisters, parents and children); and it was only at a later stage that we accumulated a sufficient number of twins to render a statistical analysis worth while. There was thus no 'inference' from twins to non-twins: on the contrary, the argument was that the conclusions drawn from studies of other relatives were still more strongly reinforced by an investigation of separated twins.

But in any case, so Stott maintains, 'separated monozygotic twins are not typical of monozygotic twins in general', and twins themselves are 'so atypical developmentally that the resemblance to a Mendelian pattern of constant expressivity that Burt discerns carries little conviction'. Stott's own explanation for the high correlations which we found is that twins are more likely than other children to have been

subject to 'gestational stress', since twinning itself 'reflects a liability to gestational stress'. Moreover, a common reason for removing a twin is the ill health of the mother, and this in turn suggests that the 'stress' was in all probability exceptionally intense during her pregnancy. Its occurrence will usually be indicated by bodily defects and ill health in the child, and, so he assures us, it would almost inevitably impair its intellectual development. That is further confirmed by the fact that twins, as a rule, are 'intellectually inferior'; the males are especially subject to such disabilities as left-handedness, defects of speech, and mental retardation generally. 'It is noteworthy', he adds, 'that Burt makes no reference to the general mental inferiority of twins or to the relative inferiority of the males, and in the present study he does not give the mean I.Q. of his twin groups'.

Stott plainly cannot have read my article with much care. The mean I.Q. for each of the three twin groups (as well as ordinary siblings) were in fact stated on p. 145 Burt, 1966). I expressly commented on 'the fact that twins, particularly uniovular, tend to have a lower average ability', and cited other investigators who had since confirmed this finding. Additional details were given in the publications cited at the end of the article. In these I discussed the probable explanations, and noted that 'left-handedness is nearly four times as common among monozygotic twins as it is in the general population' and that speech defects and mental retardation (like left-handedness) are also commoner among twin boys than among twin girls, though only to the same extent as among boys and girls of similar ages in the general population (cf. Burt & Howard, 1957; Burt, 1958*b*).

If, however, Stott wishes us to believe that the high correlations found between monozygotic twins are the result of defects caused by 'intra-uterine conditions', then it is not sufficient for him to show that such defects are commoner among twins; he must also show that the degree of the defect (i) is similar for *both* members of a pair and (ii) is itself correlated with the degree to which intelligence is impaired. No evidence whatever is adduced to support these two assumptions; and indeed, as I have pointed out elsewhere, the observations of most other investigators (as well as our own) contradict the first assumption, and include many cases conflicting with the second. Among the cases which I cited in my previous reply to Stott (Burt, 1960) I mentioned twenty-six instances of mental deficiency in an 'identical twin' which I found during investigations in L.C.C. institutions; in twenty-four of these cases the other twin was not mentally deficient. Berg & Kirman (1960) have similarly found that in forty-nine cases where the deficiency occurred in one twin, the other was *not* defective in forty-five cases. Because twins have developed within the same uterus during the same period of time, it does not follow that the pre-natal condition was the same for both, as Stott's argument assumes. The position of the two foetuses within the uterus cannot be the same; and the time, and often the mode, of delivery will be different. Further, the disturbance, whatever it is, that causes the splitting usually seems to affect the two portions of the developing embryo to different extents. As a rule, it is the weaker and less developed twin that is usually the second to be delivered, and not infrequently stillborn.

Among our separated uniovular pairs the only discernible physical defects were left-handedness (one case) and a slight feebleness in physique (two cases); and, as I pointed out, these defects occurred in only *one* member of a pair. Among the mentally

defective twins in L.C.C. institutions, both monozygotic and dizygotic, physical defects were far more conspicuous; yet they were by no means invariably common to both members of a pair, and evinced little relation to the degree in which intelligence was deficient, except in the obviously pathological cases.

Stott's references to Shields's data are difficult to understand. Shields himself (1962) thinks that physical defects may occasionally have accounted for the *differences* in the intelligence-test scores; but the notion that 'the *resemblances* might be due to pre-natal environment is', so he believes, 'improbable according to present knowledge'. Stott mentions the 'case of congenital syphilis': but Shields distinctly states that 'it was the *more intelligent twin* who had congenital syphilis and partial blindness; the co-twin (a blood donor) did not seem to have been affected'. Shields' inferences are thus the opposite of Stott's. My own conclusions were that pre-natal conditions, if anything, are more likely to diminish the resemblances than to increase them, and that their effects on the mental capacity of the twin in which they happen to occur are relatively slight.

Stott is quite wrong in supposing that I attempted to 'establish a genetic pattern from the results of twin-studies'. The 'Mendelian pattern', as he terms it, was derived from current genetic theory; and the empirical evidence which I cited as corroborating it was, not the table summarizing the results for twins, but the later table giving, for various degrees of affinity, (i) the correlations obtained in various studies by my coworkers and myself and (ii) the median correlations drawn from 129 studies by independent investigators, as compared with (iii) the value predicted by the Mendelian theory. The close agreement found for all the fifteen categories would surely constitute a most remarkable set of coincidences, were Stott's alternative explanations correct. Contrary to what Stott quite gratuitously suggests, I did *not* assume constant 'expressivity', but definitely allowed for this source of variance in the formulae used. Stott makes no reference to the table. Nor does he offer any explanation for the cases of exceptionally bright children found among the institutional cases or in families of extremely dull parents where the environmental conditions would surely have doomed every child to failure. These, as I have repeatedly argued, are well established occurrences which seem totally inexplicable on his theory.

III. THE NORMAL DISTRIBUTION

Dr Lewis begins with two points arising from my attempt to apply Mendelian principles in investigations into the 'genetical determination of intelligence'. First, he says, 'the normal distribution of intelligence defined operationally by tests such as the Stanford–Binet . . . appears as a direct consequence of the statistical criteria used to select the test items'. Secondly, he holds, 'the assumption of such a distribution, though mathematically convenient, is of little practical importance'.

He offers no formal proof that a normal distribution 'directly follows' from the statistical criteria, but merely gives a few examples showing that it *might* follow in certain cases. By selecting other examples, conforming equally well to his own principles, it is possible to produce quite a different distribution (Burt, 1957). The statistical criterion which I myself suggested (Burt, 1921) and which Lewis apparently accepts (Lewis, 1957) is that a test for age IX (say) is one that will be passed by half

the children aged 8 (and consequently by more than half the children aged 9); similarly for other ages. The data obtained show that the mental ages for the children in a given age-group (age 10, say) display a bell-shaped frequency-distribution with tails somewhat elongated, particularly towards the lower end. Now let us apply the same criterion to the measurement of height. At the time of my survey half the children aged 8 had a height of 120 cm or more; half of those aged 9 had a height of 124 cm or more, and so on. Thus, if a 10-year-old has a height between 120 and 124 cm, we can assign him a physical age of IX. The result obtained with this procedure is again a bell-shaped frequency-distribution, closely similar to that obtained with the Binet tests. Would Dr Lewis argue in the case of height the distribution found is 'not a brute fact of nature', but merely a statistical artifact resulting from the criterion used for classifying the grouped frequencies? As regards mental ability nearly every teacher of experience would agree that individuals of average ability are exceedingly common, and often, so far as general capacity is concerned, almost indistinguishable, but that the extremely dull, like the extremely able, are comparatively rare. Is this merely a teacher's illusion?

Nor can I agree that the kind of frequency distribution that actually obtains is of little practical importance. The type I have described was in fact adopted by the Education Office for London (Sir Robert Blair) as the basis for planning the number of school buildings required—'elementary', 'special m.d.', 'central', and 'secondary'. A strictly normal distribution was used by the Minister of Education when replying in Parliament to a recent inquiry about the number of highly gifted children; and in my opinion the estimate so reached was badly out. The fact is that, whether tested by the Binet scale or assessed by more accurate procedures, the distribution of intelligence does *not* conform to the normal curve. Not only are both tails (as I have just observed) much larger than a normal distribution would suggest, but the curve obtained is appreciably skewed. In terms of the old Pearson nomenclature, it is a curve of type IV, not type VII. Certainly within the limits of ± 2 S.D. the distribution can, for most purposes, be treated as approximately normal; but, when used to estimate the probable number of subnormal or supernormal individuals, it becomes decidedly misleading. As I have pointed out elsewhere, the proportion of highly gifted children with I.Q. over 175, instead of being only 3 or 4 per million as stated in the Minister's reply, appear to be 'at least 70 per million, probably more'. Mr Robb has since brought further evidence in favour of this higher estimate; and Dr Wall, Director of the National Foundation, has also expressed his agreement. But the point is of theoretical interest also. Type IV is precisely the type of distribution we should anticipate, if the genetic determination of individuals is the result of both multifactorial and unifactorial inheritance (cf. Burt, 1963 and refs.).

IV. The influence of environment

Both Stott and Lewis have suggested that the high correlation found for the separated pairs of monozygotic twins may be partly due to 'selective placement', i.e. as Stott puts it, to 'the tendency to place infants in foster-homes of comparable socio-economic level to those of the true parents'. Now for obvious reasons the majority of children boarded out come from the very lowest socio-economic groups—

those of unskilled labour, where the father is often dead or out of work, and the mother gravely handicapped by illness or other difficulties. In such cases the practice is to place the child in a *better* type of home: thus among our separated twins all, except three, who came from class VI (unskilled labour) were placed in homes of a higher class. On the other hand, it is almost impossible to find foster-parents among the professional classes who are willing to accept children from indigent widows of the same class.

Lewis, referring to Table 1 in my paper, observes that 'Burt maintains that these figures should dispose of the explanation that the high correlations are due to the way the foster parents were chosen'; but, he adds, 'in this Burt is clearly too optimistic'. To prove his point he adopts the curious method of pooling together adjacent social classes on the ground that 'the homes would be socially similar'. What similarity he can see between homes in class V (those of semi-skilled workers, who for the most had I.Q. of 90 or over and were regularly earning a decent wage) and homes in class VI (which included unskilled or casual labourers, who for the most had I.Q. of 95 or below, and were often unemployed, and not infrequently unemployable) I find it difficult to conceive. Indeed, each of these two large classes ought to be split rather than lumped together.

Possibly the naive reader, looking at Lewis's condensed rearrangement, may infer, as he does, that there was after all a positive correlation between the foster-home of one twin and the 'true home' in which the other had remained. If, however, he will take any bivariate frequency table in which there is *no correlation whatever*, and then pool the lower groups so that these are, as Lewis says, 'over 40 per cent of the cases in a single cell', he may easily be misled into fancying that he has reduced a set of uncorrelated variables to a set that are correlated. But mere inspection is always precarious. Had Lewis proceeded to calculate the correlation for his re-arranged table, he would have found that, even with his forced grouping, the correlation between the homes is actually negative!

He goes on to suggest that I should subdivide my whole sample into two main groups—children from socially similar and from socially dissimilar homes respectively. Dissimilarity is a matter of degree; and my critics would certainly have protested against my line of demarcation wherever it was drawn. What I did, therefore, was to assess the *degree* of cultural and material opportunities existing in the several homes, and correlate the differences between the children's ability with the dissimilarity between the homes in which they had been brought up. In the case of intelligence the correlation was only 0·15, a figure that is statistically non-significant. Lewis, however, argues that I ought to have taken confidence limits. These he calculates as $-0{\cdot}12$ and $+0{\cdot}40$; and he maintains that 'a population correlation near the top end of the interval is one *equally consistent* with the data'. That is surely incorrect. If we are to compare amounts of consistency, the correlation that is most 'consistent with the data' is the one which is directly deduced from the data, namely, 0·15; and of the values falling within the interval defined by Lewis those at the top end or bottom end are the least consistent. A value of $+0{\cdot}40$ would be highly significant statistically; and to declare that a highly significant value is just as consistent with the data as the non-significant value actually computed from the data seems to me to make nonsense of statistical inference. But, even if we accepted Lewis's value

of +0·40, a correlation as moderate as this still could not account for the high correlation between separated monozygotic twins that was in fact observed (0·87), particularly when the correlation between dizygotic twins brought up, not in similar, but *in the same homes*, was only 0·55.

Lewis asks how it is that the correlations for the 'final assessment' differ from those for the 'group' and 'individual' (Stanford–Binet) tests. I had thought this would be obvious from the account of our methods (cf. Burt, 1966, p. 140). We did not, as he supposes, derive the 'final assessment' by taking an average of the two test-results. The group test, which was generally administered to a whole class or even the whole school, was used partly to standardize the teachers' ratings, and partly to facilitate comparison with the results of other investigators, many of whom had relied solely on unchecked group-tests. In most cases we took the results of the individual tests as our basis. But both scores were submitted to the teachers, after the teachers themselves had made their ratings. When appreciable discrepancies appeared, the child was examined individually a third time—often after an interval, especially when he had been in poor health at the time of our previous visit. This further check is always desirable, if the best possible assessments are to be secured.

Lewis, if I understand him rightly, is anxious not so much to overthrow my conclusions as to indicate where in his view the evidence is weak, and what is required to make it more convincing. Accordingly, in his final paragraph he suggests four conditions which, if fulfilled, would in his opinion go far to appease the doubts of our critics. First, he thinks the sample of separated twins should be limited to those in dissimilar homes (i.e. I presume, in different socio-economic classes). That would mean cutting out nine cases from our sample. The remainder would then be pairs brought up in homes which culturally were very dissimilar. Secondly, the tests should be delayed until near the end of the primary school period. In point of fact, most of the tests referred to in the report were carried out at the age of 10 or over. Thirdly, he would prefer to rely on cases 'separated shortly after birth' (not, as in some of our cases, after 5 or 6 months). In five of our separated cases the mother died in childbirth; in another six the foster-homes had been found for one of the twins before the mother left hospital. For these eleven cases the correlation is 0·91, which still yields a statistically significant difference when compared with that for the dizygotic twins. But to select this batch would be rather unfair, because, if we include the next five pairs who were separated within a fortnight of birth, the difference becomes non-significant. The smallest sample that can be justifiably taken consists of the twenty-seven pairs who were separated within a month of birth. For these the correlation again rises to 0·84, and the difference is significant. This seems to meet Lewis's fourth condition, namely, that 'the correlations so obtained should be about the same level as those reported for the entire group'.

Nevertheless, it would not satisfy all my critics. Several would still argue (in the words of one of them) that 'the conditions during the first few days after birth are crucial for the healthy development of the central nervous system; and these would be the same for a given pair, but different in pairs delivered in different hospitals or at different times in the same hospital'. This objection, however, overlooks the nature of the comparison with which we are concerned namely, with unseparated dizygotic twins. Among these very much the same proportion remained in hospital for 'the

first few days after birth'. So it is quite impossible to invoke this incidental fact as the effective cause of the wide differences between the two sets of correlations.

V. CONCLUSION

The net result of these various arguments may be summed up as follows. Taking separate parts of the total evidence Stott and Lewis have certainly shown that alternative explanations would in each isolated case be *conceivable*; but they have failed to show that these alternatives are *more probable* than the explanations which I myself advanced. Unlike my own, their speculative suggestions have not been confirmed by any empirical investigation designed to test their validity, and indeed in most instances prove on closer scrutiny to be sharply at variance with facts already established by myself and by subsequent investigators. Hence, taking all the various lines of converging evidence jointly into account, I conclude that the hypothesis which I have outlined offers by far the most probable interpretation of the facts observed.

REFERENCES

BERG, J. M. & KIRMAN, B. H. (1960). The mentally defective twin. *Br. med. J.* 25 June, pp. 1911–17.

BURT, C. (1911). Experimental tests of higher mental processes and their relation to general intelligence. *J. exp. Pedag.* **1**, 93–112.

BURT, C. (1912). The inheritance of mental characteristics. *Eugen. Rev.* **4**, 168–200.

BURT, C. (1921). *Mental and Scholastic Tests.* London: Staples Press.

BURT, C. (1946). The assessment of personality. *Egypt. J. Psychol.* **2**, 65–85.

BURT, C. (1955). The evidence for the concept of intelligence. *Br. J. educ. Psychol.* **25**, 158–77.

BURT, C. (1957). The distribution of intelligence. *Br. J. Psychol.* **48**, 161–75.

BURT, C. (1958*a*). Definition and scientific method in psychology. *Br. J. statist. Psychol.* **11**, 31–69.

BURT, C. (1958*b*). *The Backward Child.* London: The University of London Press.

BURT, C. (1960). Note on Stott's article: interaction of heredity and intelligence. *Br. J. educ. Psychol.* **30**, 273–5.

BURT, C. (1963). Is intelligence distributed normally? *Br. J. statist. Psychol.* **16**, 175–90.

BURT, C. (1966). The genetic determination of differences in intelligence. *Br. J. Psychol.* **57**, 137–53.

BURT, C. & HOWARD, M. (1956). The multifactorial theory of inheritance and its application to intelligence. *Br. J. statist. Psychol.* **9**, 95–131.

BURT, C, & HOWARD, M. (1957). Heredity and intelligence: a reply. *Br. J. statist. Psychol.* **10**, 33–63.

BURT, C. & WILLIAMS, E. L. (1962). The influence of motivation on the results of intelligence tests. *Br. J. statist. Psychol.* **15**, 129–36.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (1967). *Children and Their Primary Schools.* London: H.M. Stationery Office.

GARNETT, J. C. M. (1911). General ability cleverness and purpose. *Br. J. Psychol.* **9**, 345–66.

LEWIS, D. G. (1957). The normal distribution of intelligence, a critique. *Br. J. Psychol.* **48**, 98–104.

LEWIS, D. G. (1966). Commentary on 'The genetic determination of differences in intelligence', by Cyril Burt. *Br. J. Psychol.* **57**, 431–3.

SHIELDS, J. (1962). *Monozygotic Twins.* London: Oxford University Press.

STOTT, D. H. (1959). Evidence for prenatal impairment in mentally retarded children. *Vita Humana* **2**, 125–48.

STOTT, D. H. (1966). Commentary on 'The genetic determination of differences in intelligence' by Cyril Burt. *Br. J. Psychol.* **57**, 423–9.

(*Manuscript received* 19 *April* 1967)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 163–165
Printed in Great Britain

INDISCRIMINATE OPERATIONISM: A REPLY TO KEEHN ON 'AVOIDANCE RESPONSES AS DISCRIMINATED OPERANTS'

By ROBERT BROWN
Department of Psychology, University of Exeter

Keehn (1966) makes a number of criticisms of 'traditional' studies of avoidance behaviour, and disagrees with some conclusions advanced by the writer (Brown, 1965). The writer fully agrees with the criticisms of earlier studies, and so will not consider them here. The only real point of disagreement concerns the role of pre-shock warning signals in avoidance conditioning. Keehn seems to imply in his introduction (p. 375) that the writer's views and his own are mutually exclusive; he states: '[Keehn] regards them as discriminative in the operant sense of setting the occasion for a response (Skinner, 1938); Brown considers them to be aversive.' It should be pointed out that the difference in viewpoint is not simply dichotomous; in my previous paper it was stated quite clearly that warning signals were conceived as being both aversive and discriminative, and that these functions were inseparable. On the other hand Keehn believes that a perfectly satisfactory account of avoidance behaviour can be given without attributing aversive properties to the warning signal, only discriminative ones. The purpose of the present note is, first, to show that this kind of position is completely untenable, and, secondly, to consider briefly the implications of evidence quoted by Keehn for an 'aversive' analysis of avoidance behaviour.

'Discriminative' and 'aversive' properties of warning signals

The acquisition of a discriminated avoidance response can be explained without attributing aversive properties to the warning signal, but not in the way Keehn suggests. Moreover, the explanation is *ad hoc* and intuitively discomforting. It might be argued, for example, that non-avoidance behaviour in the presence of the signal is always followed by shock more promptly than similar behaviour in the absence of the signal. Thus the signal sets the occasion when non-avoidance stimuli are (relatively) highly aversive and, by implication, when an avoidance response will receive a (relatively) large amount of reinforcement through termination of these stimuli (cf. Sidman, 1955). The intuitive discomfort arises because there is no *a priori* reason for assuming that tactile and proprioceptive, non-avoidance stimuli should acquire aversive properties when followed by shock, but that exteroceptive, warning stimuli should not.

However, in his more recent paper, Keehn (1966) is not faced with this kind of difficulty, because he attempts to explain avoidance behaviour without reference to conditioned aversive stimuli at all. It would be pleasant, in the interest of parsimony, if the concept could be dispensed with so easily; however, closer examination of Keehn's alternative position reveals some crucial weaknesses. In a strictly operational way, he argues that avoidance responses are emitted operants which can eventually come under the stimulus control of a warning signal. He is aware (*a*) that operants

are controlled by reinforcements, and (*b*) that discriminated operants require differential reinforcement. Unfortunately the argument breaks down over these two points. Concerning the first, Keehn is able to 'explain' the acquisition of an avoidance response only because he assumes that the subject can discriminate when shock is imminent and when it has been avoided. As an empirical generalization this assumption is of doubtful validity; it is well known that reliable postponing behaviour develops long before any evidence of temporal discrimination is obtained (cf. Sidman, 1966). As an explanatory statement it is clearly circular, for Keehn is trying to account for avoidance behaviour by assuming that the subject knows how and when to avoid.

He supports this argument by drawing an inappropriate analogy between food-reinforced discriminations and discriminated avoidance, claiming that 'operationally' the two situations are identical—an experimenter's button-press may deliver positive reinforcement in S^D and withhold shock in the presence of a warning signal. It may be rather an obvious point, but clearly one worth labouring, that these two operations are not comparable. The operation of 'presenting food' has an obvious and public effect on the environment, while the operation of 'withholding shock' is in itself of no meaning to the subject. It is Keehn himself who 'shifts from an operational to an intuitive stance' (Keehn, 1966, p. 378), by inferring that because an experimenter is aware that shock has been withheld, his subject should be aware too.

A final point which demonstrates the inadequacy of a 'purely discriminative' explanation comes from evidence supplied by Keehn himself. He notes that in the 'conventional' avoidance procedure, responses in the absence of the signal do not delay shock and come to extinguish, while those in the presence of the signal do delay shock and increase in frequency of occurrence. He also notes that this pattern of behaviour develops in discriminated Sidman avoidance; in this procedure, however, responses in the absence of the signal do have an effect in delaying shock. In one sense, the same reinforcement contingency is in force whether the signal is present or absent: a response cannot be followed by shock more closely in time than the specified RS interval. Thus it would seem necessary for Keehn to explain why a discrimination should develop at all, or at least why discriminated Sidman avoidance should not generate a different pattern of behaviour from discriminated 'conventional' avoidance.

'Aversive' analyses of avoidance behaviour

A theorist prepared to acknowledge the aversive properties of the warning signal might explain this fact as follows. In both procedures, the signal is the stimulus which is most consistently followed by shock, and therefore the most aversive conditioned stimulus in the situation. By comparison, behavioural non-avoidance stimuli are relatively unimportant. Hence a response in the presence of the signal receives a relatively large amount of reinforcement through signal-termination. In the absence of the signal, a response only terminates non-avoidance stimuli, thereby receiving relatively less reinforcement. Hence it is to be expected that non-signalled responses will be relatively few in number, whether they have a genuine shock postponing function or not, because of this fairly unimportant aversive role of non-avoidance stimuli.

This explanation does not solve the problem raised by Keehn: that if the signal is

aversive it should also suppress prior non-avoidance behaviour, as shock does. However, this kind of problem needs to be solved, not discarded, and the solution must be found in the laboratory, not the armchair. While the concept of a conditioned aversive stimulus has proved to be extremely useful and surprisingly resilient, the writer would be the first to agree that it needs greater refinement and specificity. It is possible, for example, that a stimulus which is just sufficiently aversive to reinforce escape is not aversive enough to suppress prior responses. But, as yet, nothing is known about the relative escape and suppression thresholds for unconditioned aversive stimuli, let alone conditioned ones. The answers to this and similar problems will come from more simply designed experiments than the ones that Keehn favours (e.g. Sidman, 1957; Sidman & Boren, 1957). It is indeed strange that Keehn should prefer experimental complexity and theoretical over-simplification when the current success of mathematical learning models suggests exactly the opposite strategy: that a complex and comprehensive theory is necessary for even the simplest of behaviour. We should not discard a concept just because it does not obey some conceptually simple rule; it would be unfortunate for psychology if the baby were always so thoughtlessly tipped out of the tub, and the bathwater retained.

References

Brown, R. (1965). Discrimination of avoidable and unavoidable shock. *Br. J. Psychol.* **56**, 275–83.

Keehn, J. D. (1966). Avoidance responses as discriminated operants. *Br. J. Psychol.* **57**, 375–80.

Sidman, M. (1955). Some properties of the warning stimulus in avoidance behaviour. *J. comp. physiol. Psychol.* **48**, 444–50.

Sidman, M. (1957). Conditioned reinforcing and aversive stimuli in an avoidance situation. *Trans. N.Y. Acad. Sci.* ser. ii, **19**, 534–44.

Sidman, M. (1966). Avoidance behavior. In *Operant Behavior: Areas of Research and Application*, ed. W. K. Honig. New York: Appleton-Century-Crofts.

Sidman, M. & Boren, J. J. (1957). The relative aversiveness of warning stimuli and shock in an avoidance situation. *J. abnorm. soc. Psychol.* **55**, 339–44.

(*Manuscript received* 16 *January* 1967)

Br. J. Psychol. (1967), 58, 1 *and* 2, *pp.* 167–195
Printed in Great Britain

PUBLICATIONS REVIEWED

The Psychology of Set. By D. N. Uznadze. New York Consultants Bureau: 1966. Pp. xvii + 251. $15.

The evidence on which this book is based comes from a series of experiments investigating the illusory perception of equal objects as unequal, following a number of presentations of unequal objects. First noted by psychophysicists in weight perception experiments, Uznadze has demonstrated the phenomenon in the various human senses, and in animals. He has shown cross-modal transfer, and studied the extinction of such illusions or sets in normal people and in psychotics.

Although numbers of subjects are often small and a variety of different responses observed, sweeping generalizations as to the nature of mental abnormality are drawn. Personality types in normal subjects are also postulated on the basis of the different characteristics of set observed in them.

Very few experiments are described fully, the author preferring to state the inferences he draws as established facts, merely alluding to the experimental data from which they are drawn. The few exceptions tend to undermine faith in these inferences. For example, one demonstration of set in animals cannot fail to succeed, as whichever of the two possible reactions the subject gives is claimed to indicate the existence of a set.

This may be so if set is, as the author claims, a necessary condition for any behaviour, even perception, to occur. The definition of set has now been generalized from illusions to an integral state in the organism. This state is supposed to arise in any situation where a need is experienced and there is an opportunity for satisfying it. This assertion is not supported very strongly by the evidence cited but probably few would reject it in such a general form.

This is very much a specialized book for the student of time errors or of Russian psychology. Others are advised to give it a wide berth, until its essence is distilled in a review article.

R. LYNN

Human Conceptual Behaviour. By Lyle E. Bourne, Jr. Boston: Allyn and Bacon. 1966. Pp. viii + 139. $2.50.

The behaviour of human beings asked to discover groupings among stimuli, usually two-dimensional geometric figures, according to rules prescribed by the experimenter has been studied recently in a number of American laboratories. The author writes (p. 45), 'Prior to 1950 there was a dearth of factual evidence on human conceptual processes; a recent upsurge of interest, however, has provided now a respectable body of knowledge', and, 'Our primary purpose is to relate the available evidence in concise and understandable fashion and not to quibble about definitional ambiguities'. Thus is cut many a Gordian knot inherited from earlier psychological and philosophical analyses of the problem.

The recent work is summarized in a clear and sensible if necessarily arbitrary scheme according to subject variables and task variables. Task variables include number of relevant and irrelevant dimensions, the role of positive and negative instances, informativeness and redundancy of cues, variation of feedback and contiguity of items. Among subject variables are memory, motivation, intelligence and transfer of training. An account of theories of conceptual thinking compares S–R association with hypothesis testing views and suggests how these may be compatible. Mathematical models of the response to cues and of the generation of hypotheses in concept attainment tasks are outlined.

Piaget's theory is deliberately omitted because, among other things, 'It diverges widely from conventional American theory'. The author acknowledges that his survey is limited by the neglect of comparative, developmental, pathological and social studies relevant to conceptual behaviour. There is little reference to work done before 1950. What the book achieves is a useful summary of recent endeavour in the field of 'concept attainment', or 'concept utilization'. There are some dangers of parochialism in time when the author writes that, 'In 1924 Heidbreder correctly anticipated two main lines of development' concerning spectator and participant

behaviour and illustrates one line with reference to a 'composite photograph' (p. 24). This does less than justice to Heidbreder, and none to Sir Francis Galton. Again he writes 'Underwood (1952) developed an important hypothesis concerning... the temporal contiguity of instances...' (p. 60). Some brief historical perspective is surely required.

MARIAN ANNETT

Psychological Tests. By EDGAR ANSTEY. London: Nelson. 1966. Pp. xix + 300. 63*s*.

Psychological Tests is something of a misnomer for this book since, apart from a few odd paragraphs, it confines itself almost exclusively to cognitive testing. Moreover a substantial proportion of the book is devoted to item analysis—in particular to Dr Anstey's own ingenious '*d* method' of item analysis. That he is aware of this top-heaviness is evident from his preface but such evidence does not justify the retention of the book's title. The chapters concerned with item analysis are highly mathematical, critical and yet constructive. Much of what Dr Anstey has to say, here and elsewhere, is well worth saying, but the book is out of balance, partly because of this odd distribution of space and partly because the author ranges from obvious statements bordering on the naïve to obscure propositions assuming a high degree of sophistication. An example of the former (with reference to construct validation): 'Confirmation can never be complete, but the construct is strengthened by any piece of positive evidence and weakened by any piece of negative evidence.' An example of the latter: 'As $N \to \infty$ the distributions of $m(R)$ and $m(W)$ for different samples tend to normality, and in the limit d is distributed normally about a mean $[\mu(R) - \mu(W)]$ with variance $\sigma^2(R)/R + \sigma^2(W)/W$. On the assumption that $\sigma(R = \sigma(W) = \sigma$, the variance of d reduces to

$$\sigma^2\left(\frac{1}{R}+\frac{1}{W}\right) = \frac{\sigma^2 N}{RW} = \frac{\sigma^2}{Npq}.'$$

This is characteristic of a lot of the writing and it is unlikely that it will prove intelligible to all the 'people who are concerned with intelligence and other psychological tests in that they have to interpret test results or to decide whether or not tests should be used in a particular context.' Yet these are among the readers, so the preface explains, for whom the book is intended.

The chapters on item analysis are valuable although, in the reviewer's opinion, Dr Anstey sometimes sacrifices psychological understanding on the altar of statistical rigour—despite his repeated statement of the need 'to combine objective information about items with psychological insight.' In his interesting chapter on 'Doubtful test items', for instance, he appears to make little or no distinction between poor questions, which should never have reached the stage of item analysis, and questions which his subjects have answered poorly and indefensibly (e.g. 'overlooking the fact that the first term [in a given numerical series] did not fit in with the rest'.)

Other topics include the meaning of 'intelligence', job analysis, instructions for devising a new test, some discussion of validation methods, a chapter on tests for the highly intelligent and a short piece on the effects of practice and coaching. Every chapter ends with a useful summary. At the end of the book is a summary of conclusions, a bibliography of some 120 titles, an index of tests discussed (incomplete) and appendices which include tables for *r* and *phi* and item analyses. The discussion of test validation is brief and inadequate, skating rather lightly over some of the inherent difficulties. On the other hand, the chapter on high-grade tests contains some valuable suggestions and caveats.

Dr Anstey's main weakness is perhaps a tendency to oversimplify: he accepts too readily the notion that there is such a thing as a 'true score'; that the distinction between 'speed tests' and 'power tests' is tenable; that the presentation of preliminary examples, which are already completed, is sufficient to ensure that justice is done, and seen to be done, by the subjects; and that 'the validity of a test can be defined as "the extent to which the test measures what it is required to measure".' His main strength (linked no doubt with this weakness) is a gift for clear exposition. This makes the book a great deal more readable than many others on the subject. There are a few misprints and factual errors—including one in a sample question—which it is to be hoped will be corrected in future editions.

A. W. HEIM

Personality, Social Class and Delinquency. By J. J. CONGER and W. C. MILLER. London: Wiley. 1966. Pp. viii + 249. 60*s*.

This book is an important addition to the growing literature of investigations into the psychosocial characteristics of boys who become officially classified as 'delinquent'. Earlier studies in this field, such as those of Healy and Bronner and the Gluecks have often suffered from the error of hindsight: the investigators were generally aware of which individuals had in fact become officially 'delinquent' before they made investigations into their personal characteristics and family backgrounds. The present study does not suffer from this drawback, since data about the experimental group and the controls were being collected from an early age before any delinquency took place.

This investigation was part of a much larger one involving over 2000 schoolboys of 15 years of age. From this population all those who were later in trouble with the juvenile court by the age of 18 years were identified, and matched with non-delinquent controls according to personal and social criteria. In this matching process nearly 25 % of the delinquent groups were dropped, because adequate matches could not be found for them. This point does not seem to have worried the authors, but it does stand out as a serious flaw in their study, for it is impossible to say just what were the characteristics of this special group.

The most valuable measures that were available came from the teachers' ratings on these boys which had been periodically recorded since their kindergarten years. The usual overall interrelations held good that boys noted as troublesome at an early age by their teachers later come into conflict with the law. This truism, however, was contradicted in part by more sophisticated methods of analysis than have usually been employed in such studies. By partialling out, it was possible to show that in certain sub-groups, defined by criteria of intelligence and socioeconomic status, the expected relations either did not exist, or were indeed actually reversed. This is an important advance in methodology which may yet dispel some of the time-honoured myths about delinquency which have been respected for too long.

As well as teachers' ratings, the research used the results of group testing the major population. Some of the tests were of the multi-scale variety, such as the Guildford–Zimmerman Temperament Survey, and it might be suggested that in spite of the wealth of names given to the many sub-scales, the results of this questionnaire material might be expressed in terms of fewer and simpler personality factors. These results were not as meaningful as those derived from the teachers' ratings, nor were those obtained from a modified Thematic Apperception test.

Perhaps in this study, as in many others, the boys who were least cooperative and academically able, (who are often most likely to be delinquent) may have partially frustrated the personality-testers by means of their very deficiencies.

H. B. GIBSON

Psychological Principles appropriate to Social and Clinical Problems. By J. C. RAVEN. London: H. K. Lewis. 1966. Pp. 187. 25*s*.

This book, by a founder of British clinical psychology, is in two parts, Practical problems and Theoretical principles. The first part has chapters on the interview, comparative matching, questionnaires, and intelligence testing. The second part considers psychological theory, attention, intentional behaviour, value, orderly thinking, individuality, and personality. Hostility might be provoked by the rather few fashionable references. Freud, Hull, Jung, Watson, receive one each—no more than J. W. Dunne, Hume, Max Scheler and Titchener.

This book is in fact, a 'highly individual production. It expounds (the author's) own approach to psychology and indicates some of the practical measures which this theoretical structure appears to indicate'. The chapter on intelligence testing, written by an authority, repays close reading. Raven's *Progressive Matrices* and vocabulary tests, although internationally known, do not receive, perhaps, the value judgement they merit. Often, they are regarded as group tests, inferior to polymorphous tests, though there seems little doubt where the theoretical and factorial purity lies. The questionnaire chapter is really a particular research report, although this does raise general criticisms of inventories and what they are alleged to measure clinically. There are important chapters on interviewing and the technique of comparative matching. The latter technique is an excellent way of 'counting heads', a necessary but often neglected priority

in personality studies. It is a pity that Foulds' TAT work is not included, showing how this technique can be used to draw up measuring scales.

In the book's second section, the writer's theoretical opinions are given. Contrasting viewpoints are interestingly compared so that the limitations of particular standpoints such as the behaviourist, the psycho-analytical, and the physiological are exposed. Some of the judgements may seem a little old fashioned—but what's new about behaviourism and psychoanalysis? Yet Raven's own viewpoint, a sympathetic 'middle of the road' one, is not clearly demonstrated by application, to be more useful. For example, a psychosis is defined as 'a mental illness characterized by some degree of personality disintegration'. Schizophrenics 'experience difficulties in establishing satisfactory boundaries between themselves and other people'. Such hypotheses seem reasonable and are similar to those used by other workers. But such hypotheses remain difficult to validate. We still need to see, as well as believe, that the majority of schizophrenic patients show *x* and the majority of non-schizophrenic patients *non-x*. Much in fact, of what the writer says about various disturbances seems unexceptional but only, perhaps, because it is a polished distillate of classical descriptions. The clinical psychologist has the difficult day-to-day task of trying to make classical traditions work. Such difficulties, in the personality field, must indicate that traditional beliefs are either imprecise or invalid, despite their common appeal.

This book is an interesting but mixed bag, attempting as it does, to tackle the invidious problem of presenting an all round account of the individual. It will not be appreciated by the more inflexible adherent to a particular point of view, but the more adventurous reader might find much to stimulate him.

J. E. ORME

Toward a Field Theory of Behavior: Personality and Social Structure. By J. MILTON YINGER. New York: McGraw-Hill. 1965. Pp. ix + 365. 72*s*.

Theories in Social Psychology. By MORTON DEUTSCH and ROBERT KRAUSS. London: Basic Books. 1966. Pp. x + 244. 25*s*.

Small Group Research: a Synthesis and Critique of the Field. By JOSEPH E. MCGRATH and IRWIN ALTMAN. New York: Holt, Rinehart and Winston. 1966. Pp. ix + 601. $12.50.

The pendulum has swung again and, after some four of five decades of data-gathering, it is again becoming acceptable in social psychology to venture on theory construction. Three recent books, in quite different ways, make a contribution to this thinking about theory. The first is concerned with theory at an all-embracing level and is written by someone whose own discipline is not psychology; the second is a review largely of theories of the middle range and is written by one who has himself made a substantial contribution both to experimental and theoretical social psychology; and the third concentrates on one area within social psychology, an area in which the two authors are practitioners.

Yinger's aim, if achieved, could have the effect of making conceptual sense of the uneasy position social psychology now occupies between psychology and sociology. Himself a sociologist and anthropologist, he endeavours to develop a perspective which would help people studying many aspects of human behaviour to see better the wider ramifications and implications of their own work. The programmatic statements of the first part of the book (for example, defining the unit of analysis as 'the field in which the individual and the social structure meet'), although beyond reproach in their reasoning, are unlikely to move psychologists greatly because of the familiar and hollow ring that the main concepts have. The second part, in which he discusses more fully the elemental constructs of the theory and their interrelations, is more persuasive, partly because he 'touches down' from time to time with empirical studies. It is in the third part, in which analyses are presented of topics like anomie and alienation, delinquency, prejudice and discrimination, that one has most opportunity to admire the searching tool that a 'field' perspective appears to be. But of course the truth may be simply that Prof. Yinger is gifted at seeing all round problems. Without a greater formalization of the general model it is difficult to be sure that others could learn to use it as effectively. However, this book (particularly if it should be issued in a cheaper edition) could conceivably be material for a stimulating seminar on how to think about problems in social science.

The main sections in the review of theories made by Deutsch and Krauss resemble rather closely the early chapter headings of Lindzey's *Handbook of Social Psychology*, now twelve years old: gestalt psychology, field theory, reinforcement theory, psychoanalytic theory and role theory, preceded here by a discussion of the function of theory and followed by a review of trends in social psychology. The content gives full representation to the theoretical preoccupations of the intervening twelve years, and one wonders how this line-up of theoretical protagonists remains relevant. In practice, however, it proves illuminating to encounter modern middle range theories in the context of their more distant ancestry; to meet cognitive dissonance as a field theory and to find Skinner and Homans cheek by jowl in the section on reinforcement theory. This book, like the whole *Basic Topics in Psychology* series, is intended mainly for the instruction of students and is unlikely to fail in its aim.

Deutsch and Krauss, in their introductory chapter, advance the rejigging of its definition that social psychology has been receiving of late: 'Social psychology is in its infancy. It has only begun to identify a distinctive subject-matter relating to human interaction' (p. 12). Several decades of experimentation in small group laboratories have made a sizeable contribution to this re-definition. The proliferation of small group studies has presented a challenge to collators which has been taken up by several in recent years. McGrath and Altman have set about the task, not by superimposing categories of their own, but by reference to a detailed operational analysis of the variables studied. They regard their work, therefore, as in part a contribution to the science of the classification of information, and in part 1 the small-group field is treated simply as a case study for demonstrating the principles of classification they are adopting. An interesting feature of this is that it has an inbuilt validation system: by classifying with reference to the main variables of experiments they can check whether the relationships among variables established by the experiments provide a quantitative verification of the logical relations. This exercise is carried out on a sample of about 250 studies, selected from over 2000 presented in the final bibliography. Part 3 summarizes the variable classes which have emerged from this analysis and the empirical findings about the relations among them which the 250-odd studies have produced. (Unfortunately there is no index of variables. This would not only make the whole thing more useful as a work of reference but could possibly have eliminated the necessity to present every relationship twice over.) Part 3 concludes with a complete annotated index of the 250 studies.

It will already be apparent that this is a useful reference book on small-group literature. Its value is enhanced by the second part, in which the authors present what they call perspectives on the small-group field—including a description of its 'culture'. Particularly useful here is a chapter in which generalizations from the data are summarized. Although these are often negative in form, they do suggest where further research effort needs to be directed. The authors also devote a chapter to criticism of general methodological issues, noting particularly the lack of replication of experiments, the need for a common language, for cumulation of knowledge, for construct validation, for 'programmed diversity of research settings' and for multivariate approaches. Finally they join the small chorus calling out for more theory. VERELMA VENESS

Principles of Psychopathology. An Experimental Approach. By BRENDAN A. MAHER. New York: McGraw-Hill. 1966. Pp. 525. 72*s*.

To the casual observer the title of this book may belie its contents, but its subtitle and a quick glance through its pages soon dispel any notion that this is yet another exemplification of psychoanalytic principles with which, in the minds of most of us, the term 'psychopathology' has become associated. From the outset Prof. Maher makes it very clear that what he means by psychopathology entails the application of strict experimental method to the study of abnormal behaviour. This use of the term disappeared somewhere around the turn of the century when psychiatry and experimental psychology parted company. Its reappearance in the title of this book is a sign that the experimentalist and clinician are beginning to make friends again, and that the accumulated evidence in experimental psychopathology is now respectable and coherent enough for it to be brought together in an introductory text for students, at whom the book is mainly aimed.

The first five chapters are devoted to methodology and a discussion of some basic concepts in learning theory and psychophysiology which experimental psychopathologists have found

useful in their analysis of disordered behaviour. Subsequent chapters deal with the application of these concepts to the study of neurosis, psychosomatic disease, brain damage and functional psychosis. Two final chapters cover psychological and physical methods of treatment.

The chapters on neurosis and character disorder give an adequate account of some of the experimental work on anxiety and conflict, although the selection of evidence is somewhat unsystematic and there is a failure to relate these problems to a broader theoretical framework of personality structure. In this connexion there are some rather too obvious omissions and one comes with surprise upon statements like 'No hypothesis purely in learning theory terms has yet been advanced [*sic*] to account for the acquisition of psychopathic behaviour.' Perhaps Prof. Maher feels that Eysenck's views on the subject have been well enough advertised already.

The best part of the book is certainly the coverage of schizophrenia, to which the author devotes four chapters dealing with clinical, biological, genetic and behavioural aspects. These chapters will be a godsend to students trying to find their way through the bewildering and often contradictory evidence on schizophrenia. It is a pity, though, that little attention is paid to affective disorders, especially the problem of depression which, surprisingly, rates only a one-page reference in the index. Surely interest in depression has generated enough experimental work to warrant its inclusion in any discussion of the functional psychoses.

For those of us who teach clinical and abnormal psychology from an experimental point of view the book will help to resolve the dilemma we have all faced when our students have asked us to recommend a good introductory text. Even for postgraduates Eysenck's *Handbook of Abnormal Psychology* is hard going and what has been needed for some time is a 'Munn' of abnormal psychology. Prof. Maher has gone some of the way towards providing it.

GORDON S. CLARIDGE

Eliminating the Unconscious. By T. R. MILES. Oxford: Pergamon Press. 1966. Pp. xviiii + 171. 17*s*. 6*d*.

This interesting and suggestive book constitutes, in the author's words, 'a plea that in discussing the findings of psychoanalysts we should *talk differently* and in particular *should abandon all language suggestive of traditional philosophical dualism*'. In terms of the subtitle of the book, Miles proposes 'a behaviourist view of psychoanalysis'; as he himself admits, this title was deliberately chosen to be provocative. By behaviourism, it is important to note, Miles does not refer to the metaphysical or the methodological versions distinguished by Mace, but to the analytical, i.e. he postulates that 'statements about mind and consciousness just turn out to be, on analysis, statements about the behaviour of material things'. This attitude is seldom combined with a favourable view of psychoanalysis but Miles makes out a good case for the lack of any logical contradiction. The reader may already have noted a somewhat philosophical odour, and indeed much of the book is concerned with philosophical issues rather than with factual ones. Miles is not concerned, as were, for instance, Miller and Dollard, with interpreting the alleged facts and laws of psychoanalysis in terms of the propositions of modern learning theorists; indeed he makes no reference to these and similar attempts to reconcile psychoanalysis and behaviourism. It would have been interesting had he done so.

As it is, the book is difficult to evaluate or even discuss. What Miles is trying to say is that those people who declare psychoanalysis to be outside the scientific pale are wrong, and he gives a lengthy discussion of the point. Unfortunately he follows precedent in not discussing important objections made by acknowledged experts; there is no mention, for instance, of Popper and his argument that psychoanalysis is not scientific because its statements are of a kind which cannot be falsified. Instead Miles sets up a man of straw whom he then easily demolishes; this is not a very convincing exercise. Indeed Miles seems to realize that he is batting on a poor wicket, and he ends up by concluding, not that psychoanalysis is a science, but rather that in principle there is nothing to prevent it from being a science if only it were different to what in fact it is. Whether this is a very meaningful statement, or a very helpful one, does not seem a very important question; if psychoanalysis were to change so much in the direction desired by Miles it might become a science, but it would certainly cease to be psychoanalysis. This book is indeed provocative and well worth reading even though the reader may not agree with all the points made by the author.

H. J. EYSENCK

Motivated Learning: A Developmental Study from Birth to the Senium. By MIRIAM E. HEBRON. London: Methuen. 1966. Pp. 264. 35*s*.

This book attempts a highly complex task—to relate theories of motivation and learning to various stages of development from birth to old age. Piaget's system is used to provide a general framework, supplemented by reference to the works of Freud, Pavlov, Hull, Lewin, Mowrer and Luria, amongst others. No specific introduction to theories of learning and motivation is given and maturational factors are only considered incidentally. In so far as one is able to separate 'motivation' from 'learning', the former seems to receive rather more attention than the latter in spite of the title. However, the importance of cognitive experience in the course of development from the sensori-motor stage onwards is continuously stressed.

An interesting first chapter on the physiology of motivation and learning introduces the study, and throughout the text possible physiological correlates to psychological events are offered. Just under one half of the book is devoted to child development, then follow two chapters on adolescence and the book concludes with another two on maturity and old age.

Inevitably perhaps, in combining studies in motivation and learning, the subject-matter ranges far and wide, covering such topics as attention, parental influence, the growth of self-awareness, thinking and social skills. It is impossible within the space of 233 pages to cover such a multiplicity of topics fully and students may find themselves confused by the very extensiveness and complexity of the information presented. This is accentuated by a highly condensed style of presentation phrased in language which sometimes seems unnecessarily elaborate. Such factors may produce difficulties for those who have not read much contemporary psychology. In a book of this kind further amplification throughout the text would have been helpful. The absence of a concluding chapter summarizing the writer's views is particularly unfortunate when the author has surveyed such an extensive field and collated so much interesting material.

The book is described by the publishers as directed towards 'the more advanced student preparing for the newly instituted degrees in education, to postgraduates in education, and to students in their third year of degree courses in psychology and sociology'. The relevance to education of some of the material presented is self-evident, but it is doubtful if many students in education or sociology would have sufficient knowledge of psychology and physiology to appraise the contents adequately.

There is a great need for more British text-books in developmental psychology. This is a contribution which third-year psychology students could read with advantage, but perhaps after tackling such difficult material the author would consider writing a further text with a wider appeal.

R. MALIPHANT

Contrary Imaginations: A Psychological Study of the English Schoolboy. By LIAM HUDSON. London: Methuen. 1966. Pp. 181. 25*s*.

This is an unusually readable and versatile book, full of illuminating ideas about originality in the academic or scholastic world, and of suggestive speculations about the motives which lead men and women to take up creative work in the arts and the sciences. Its pages deserve the close attention of every psychologist, educationist and teacher. Unfortunately the particular conclusions which have hitherto chiefly attracted notice are certain incidental generalizations based largely on misconception and often on fallacious argument; and it is important that readers who are not versed in the copious literature of the subject should be aware of their questionable nature.

Dr Hudson begins by confessing 'a distrust of the statistical approach to complex problems', and launches a vigorous attack on 'mental testers' both in this country and America who have relied on 'factor analytic studies'. Their 'central failing [we are told] is the neglect of inconvenient evidence'; they 'display no special knowledge of books, painting, music . . . or other aspects of their culture'; and 'the men who founded [such techniques] were not interested in individual human beings'. This is surely quite untrue. Galton and Binet are full of intensive studies of 'individual human beings'. Later books, based on mental testing and factor analysis, freely illustrate and supplement their statistical results by case-histories quite as detailed as those reported by Dr Hudson himself. Both psychological tests and factorial methods have been extended into the fields of literature, painting, music and other 'aspects of culture'.

The section most frequently quoted is that entitled 'The downfall of the intelligence test'. Here Dr Hudson attacks those who use 'the I.Q. tests', for their belief that 'armed with such tests the psychologist can probe the innermost recesses of our minds, and predict our future accomplishments unerringly'. Now these are precisely the misconceptions which 'intelligence testers' have been at pains to repudiate from the very outset. All that they have claimed is that intelligence tests will provide, within a certain margin of error, a helpful estimate of a child's innate general ability; from the days of Galton onwards they have always insisted that a man's accomplishments depend on many other factors—his special abilities, his dominant interests, his temperamental and moral qualities, and his physical health. Nor have they ever suggested that their predictions were 'unerring'. On the contrary, the statistical device of correlation was expressly adopted to measure the degree of error affecting tests of different types. And when Dr Hudson goes on to assure us that 'psychologists have neglected to collect the evidence that would prove their suggestion false', and that 'it has taken 50 years to discover what could have been ascertained overnight', he seems to be merely demonstrating (what he frankly acknowledges on another page) his 'ignorance of the literature'.

His own evidence is limited to results obtained with a single type of test applied to highly selected groups—'reasonably clever boys' of 15–17 in the upper forms of grammar and public schools. He finds that 'the only significant difference' revealed by the test he has used is the distinction between 'those who go to a university and those who do not'; it does not distinguish between scholars or exhibitioners at Oxford and successful entrants at other universities. On the other hand, the 'verbal or numerical' and the 'diagrammatic' items in his test do display a correspondence with the 'distinction between the arts and the sciences'. But none of this justifies a sweeping condemnation of intelligence tests appropriately applied and interpreted. All that it suggests is (i) that the particular intelligence test which Dr Hudson has chosen was not very suitable for the purpose in view; (ii) that the correlations to be obtained from groups already selected for general ability are likely to be low; and (iii) that, as the age of the pupils increases, the importance of individual differences in general ability decreases markedly when compared with individual differences in special abilities and special interests; each of these three points has already been amply demonstrated by 'mental testers' in numerous articles that Dr Hudson has apparently overlooked.

The greater portion of the book, however, is concerned with two contrary types of 'imagination', or rather of thinking—the 'convergent' and the 'divergent', and with the way this difference seems to correspond with an interest in science or in arts. 'This discovery', says Dr Hudson, 'is central to the argument in this book'. It is by no means so new as he supposes. Factor analysts long ago discovered a bipolar factor turning on an analogous distinction and variously described as 'fixating' or 'fluctuating', 'analytic' or 'synthetic', 'demonstrative' or 'productive' types of thinking; and at the same time it was shown that these contrasted tendencies were correlated with a bias towards scientific or mathematical and literary, artistic, or informational subjects respectively, in short, towards one or other of the 'two cultures'. Dr Hudson's outstanding merit is to have thrown so many gleams of light on this familiar distinction—on its nature, its origin, and its general effects.

According to Dr Hudson the 'converger' is one who does best at problems with only a single correct answer, and the 'diverger' one who does best at 'open-ended' problems (i.e. those which require what Stout called 'divergent' associations, though Stout is not mentioned). Since Dr Hudson supposes that intelligence tests 'have one assumption in common—that each question has only one right answer', he identifies the 'converger' with the person with a 'high I.Q.'. In point of fact earlier tests of intelligence contained many open-ended questions: the Binet scale has more than half-a-dozen; so has the original analogies test which Dr Hudson cites. The restriction to questions with only one right answer was a device introduced merely to simplify marking; and certainly among those with high intelligence as ordinarily defined there are as many divergers as convergers. But these are minor criticisms.

The main value of the book lies in the detailed discussion of the intricate processes, temperamental as well as intellectual, which underlie original or creative thinking. I doubt whether his descriptions would apply to all who show originality or inventiveness; but they certainly hold good of many. And throughout the chapters his points are illustrated with a fascinating wealth of detail. Above all, let us be grateful for a writer who can express himself with such clarity of style and convey a complex argument in words so free from needless technicality and jargon.

CYRIL BURT

Studies of Troublesome Children. By D. H. Stott. London: Tavistock. Pp. ix + 208. 32*s*.

This book consists of three parts, the first and longest of which was published two years ago under the title *Thirty-three Troublesome Children*. It is based on the 1964 Convocation Lecture of the National Children's Home in which Dr Stott gave an account of a recently completed study. In 1961 he 'was granted facilities by the Corporation of Glasgow Education Committee to carry out a small experiment in the detection and prevention of delinquency in co-operation with their school welfare officers'.

From a complete year's entry into twelve secondary schools in high-delinquency areas, composed of 1034 12- to 13-year-olds, some 109 children appeared to be 'delinquency prone' according to the results from the author's Delinquency Prediction Instrument. Teachers were then asked to complete the full Bristol Social Adjustment Guides (also developed by Stott) for this group of children from which a 'delinquency-prediction score' was obtained. On the basis of a case conference between the author and a team of specially trained school welfare officers thirty-three children were chosen for 'situational treatment', which included co-ordinating the efforts of other social agencies.

For each of these children, brief case-studies are presented which are claimed to show that 'probably as many as 80 % of the sample could be described as neurally impaired'. It might be questioned whether such a blanket term as 'neurally impaired', which appears to include gross brain damage on the one hand and subtle neurological deficiencies on the other hand, is of much practical or even theoretical value; nor are diagnostic criteria of neural dysfunction as precisely defined as would seem desirable considering its central importance in Stott's hypothesis of congenital temperamental vulnerability.

The last two chapters of part 1, respectively entitled 'Pointers to treatment' and 'Remedies', are concerned with practical issues. In the former, the responses of the thirty-three children, to the various forms of situational treatment devised for them are briefly described. The second chapter sets out 'A programme of practical measures' for dealing with the problems of behaviour-disturbance and delinquency. Among the eight listed, Stott considers the establishment of a school-based service of trained welfare officers to be of central importance.

In part 2, entitled 'Truants', the author gives an account of a validation study of his Delinquency Prediction Instrument which he began in 1963 at the invitation of UNESCO. Again school welfare officers collaborated and their Association undertook the detailed administration of the whole project. Each officer selected his two most recent truants and a sample consisted of 305 children chosen by 168 welfare officers all over Britain. For each truant two controls were obtained from within the same school, the criteria for matching being sex and being born on the same day or the day nearest to it. The second control was chosen in order to obtain a measure of cultural influences. Here, in addition to sex and year-group, the criterion was that the control should live in the same street, or as near as could be found, as the truant. Class teachers then completed Bristol Social Adjustment Guides for the truants, for the birth-date and for the neighbourhood controls. Education welfare officers filled in a schedule for each truant which covered his home background as well as the frequency and reasons for the truancy.

The mean score for the truants was found to be three times that of the neighbourhood controls as well as that of the birth-date controls. Some 63 per cent of the truants were maladjusted according to the Guides' classification, compared with only 10·4 per cent of the controls as a whole. Stott concludes that truancy is in the main a symptom of 'fairly severe maladjustment among a class of temperamentally vulnerable children exposed to a degree of stress which exceeded their emotional reserves'; he also states that his findings show that 'the Bristol Guides are reasonably free of cultural and sex bias'. A rather strange claim, perhaps, for instruments specifically aimed at assessing social adjustment? Again it is stressed that properly qualified school welfare officers could play a central role in the prevention and treatment of truancy and problems related to it.

Stott's emphasis on the potential contribution of the school in the early detection of deviant behaviour is widely accepted. Whether a convincing case has been made for school welfare officers to undertake the main burden of the task, both in relation to assessing the home background and with regard to the 'situational treatment' of the child and his family, is open to doubt. Moreover, Stott has little to say about the role of the class-teacher, the head-teacher and

indeed the school as a social system and thus as a potentially major, constructive influence both on its pupils and on the whole community in which it functions.

Last but not least, some doubt will remain about the validity of both the Bristol Guides and the more recently developed Delinquency Prediction Instrument until reliable norms, preferably based and standardized on a national sample, become available. However, none of these considerations invalidate the basis of the author's practical proposals. These rest on the argument that the incidence of maladjustment among the school population is likely to be 10 per cent (and there is now a fair amount of evidence to indicate that this may well be the case); and that existing school psychological and child guidance services are numerically quite inadequate to deal with a problem of this magnitude. Hence the need to consider alternative methods and services.

The third part of the book, entitled 'Towards an understanding of behaviour disturbance', is the most ambitious and wide-ranging. In the first chapter, 'A law of multiple impairment', Stott summarizes his own work over the past 10 years on which he bases the promulgation of this law. The next chapter, 'Types of maladjustment', outlines the procedures he adopted to arrive at his classification of syndromes of maladjustment. In the final one, the author attempts to outline (in a brief 11 pages) 'A general theory of human motivation'. Not unexpectedly, the attempt is disappointing: too many hypotheses are put forward; supporting arguments are too brief and sketchy; and too many propositions are inadequately conceptualized. This applies, too, to his central thesis that the majority of children, showing behaviour disturbance, also suffer from some 'congenital neural impairment or dysfunction'. Indeed, this highly condensed last chapter is not easy to follow and at times the author seems to put forward contradictory theories of motivation and explanatory models of behaviour deviation. Thus while the empirical findings and practical recommendations presented in this book are a valuable contribution to the study of delinquent and otherwise disturbed children, Stott's theory of human motivation and of abnormal behaviour are at best suggestive and at worst idiosyncratically speculative.

M. L. KELLMER PRINGLE

Cultural Factors in Delinquency. Edited by T. C. N. GIBBONS and R. H. AHRENFELDT. London: Tavistock. 1966. Pp. xii + 201. 35*s*.

The World Federation of Mental Health, under the chairmanship of Dr Otto Klineberg held an interdisciplinary conference at Topeka in order to study cross-cultural factors of delinquency. The editors, both English psychiatrists, have diligently converted the papers given at that conference into most readable little book which is full of useful information about such topics as the rise of crime rates in most countries during their varying stages of industrialization in relation to various well-known theories. The book falls into two parts and deals with such topics as definitions, comparability of statistics, learning theory and research. Few chapters are tackled in any depth; there are many omissions, especially on research, and there is an arbitrary selection of works. The volume falls between two stools and is likely to whet the appetite of psychoanalysts on the one hand and social experimental psychologists on the other without fully satisfying either. Very little of the 'hard' approach that psychologists are used to is to be found in this book, even in this under-researched field. This may partly be because of the construction of the conference, which consisted mainly of psychiatrists and sociologists and was notable for the absence of experimental or clinical psychologists. Nevertheless, the book contains some useful information, and will be of value to criminologists on this account.

R. G. ANDRY

Depressive States: A Pharmacotherapeutic Study. By A. HORDEN, C. G. BURT and N. F. HOLT. Springfield, Illinois: Charles C. Thomas. 1965. Pp. xvi + 166. 60*s*.

This is, in many ways, an admirable investigation of the relative effects of imiprine (tofranil) and amitriptyline (tryptanol) upon depressive states in women. Two kinds of measures were mainly used: (i) the averaged independent 'blind' responses of two doctors on the Hamilton Rating Scale, based on a joint interview; and (ii) the clinical decision to discharge the patient, taken either four or six weeks after the beginning of treatment. There were 69 subjects in the tryptanol and 68 subjects in the tofranil group. The groups were matched on all the relevant variables; subjects were randomly assigned to one or other of the two treatments.

By the end of six weeks, 81 per cent of the tryptanol group and 54 per cent of the tofranil group were discharged ($P = 0{\cdot}002$). A six-month follow-up showed that 84 per cent of the tryptanol discharged and 79 per cent of the tofranil discharged were still well.

There seemed to be little difference between the two drugs for patients between 30 and 39 years of age; 17 out of 22 were discharged in the tryptanol group and 15 out of 23 were discharged in the tofranil group. There also seemed to be little difference between the two drugs for reactive as compared with endogenous depressives; among the reactive depressives, 23 out of 27 were discharged on tryptanol and 22 out of 30 on tofranil.

The study presents a number of problems. First of all, the apparently clear overall superiority of tryptanol is obscured by the fact that deluded patients, whose prognosis is worse, were more numerous in the tofranil than in the tryptanol group. When the deluded patients are excluded, 88 per cent of the 52 remaining tryptanol and 73 per cent of the remaining 37 tofranil patients were discharged. (Reviewer's $P < 0{\cdot}05$ on one-tail test.)

A second problem is presented by an apparent contradiction in one outcome of two different methods of analysis. The analysis of the relationship of age to discharge indicated that the superiority of tryptanol was not clearly demonstrated in patients under 39 years of age. However, an analysis of variance carried out on percentage improvement scores on the Hamilton rating scale did not result in a significant interaction between age and treatment; the main effect for treatment was still highly significant. Difficulty is increased by the fact that a correlation of 0·85 is reported between the discharge rate and percentage improvement scores on ratings. The reviewer may have overlooked some obvious fact.

A third problem is presented by the fact that the initial selection and the final discharge of the patients for the study was not in terms of some measurable aspect of the patient's behaviour or talk. For example, selection was in terms of an entirely uncontrolled method; on the basis of somebody apparently having judged that the primary clinical manifestation was 'a persistent alteration of mood exceeding customary sadness' (p. 47). This kind of difficulty pervades most clinical psychological research, including that of the reviewer himself. We cannot ensure the reproducibility of our results until we overcome this difficulty.

The authors make numerous additional observations on the nature and treatment of depressive states, such as the relative lack of correlation between depression and anxiety, and the change in severity of certain symptoms with age. They also link the whole of their research with a consideration of the relevant literature. The writing is clear. Altogether, this is a useful book, both for research workers and students.

M. B. Shapiro

New Directions in Psychology, III. By G. Mandler, P. Mussen, N. Kogan and M. A. Wallach. N.Y.: Holt, Rinehart and Winston. 1967. Pp. ix+289. 38*s*.

The three contributions included in this book, like those in *New Directions* I and II, are intended to present new developments in specialized fields of psychological research in a form suitable both for the student and for the informed layman. They aim to give an up-to-the-minute account of the sort of questions currently claiming the attention of research workers, with enough of the background to enable the intelligent reader to grasp the problems involved. The three topics included in this volume are not otherwise related to one another in any obvious way, and must therefore be reviewed independently.

'Verbal Learning', by George Mandler

The title of the first contribution perhaps needs some explanation. The author is not basically concerned either with how young children learn the meaning of words, or with how older children acquire concepts through verbal instruction. Instead, his interest lies in the formation of principles to explain how linguistically competent adults perform in word-association experiments in which two or more words are presented in temporal proximity and the subject is required to learn to connect them, despite the fact that they are not obviously related to one another in any meaningful way. Perhaps the oddest thing about this research problem is that it is possible to trace a continuous interest in it back through Thorndike to Ebbinghaus, and that it is still regarded as an important area of psychological interest, despite the fact that no one has

successfully formulated any really adequate general principles to account for the way in which such arbitrary associations are formed. What seems to have happened is that successive generations of psychologists have each managed to apply the currently fashionable concepts and jargon to the same hoary problem; and one begins to suspect that the problem itself has remained popular more because it lends itself to neat experimental manipulations than because it has been proved to have any fundamental theoretical significance. The conclusion which finally emerges is that it is exceedingly unlikely that any one principle, or set of principles, will account for the ease with which particular individuals learn particular associations, and that, in reluctant response to this uncomfortable fact, interest has now shifted to an altogether different problem: namely, how people invent and formulate rules and strategies to cope with any kind of learning problem. For the future, this new direction may well prove scientifically fruitful: at present it seems to be an area in which psychology is still struggling to find meaningful questions which could be answered, even in principle, by empirical studies.

'Early Socialization: Learning and Identification', by PAUL MUSSEN

This second contribution could hardly be more obviously separated both in scope and character from the previous one. Mussen is concerned with the very broad problem of how the character of the human child is moulded through the process of social interaction with his parent. The theoretical context in which Mussen discusses the relevant research poses the recurrent problem: are the principles of reward-learning sufficient to account for the way children gradually acquire moral inhibitions and standards from their parents, or is it necessary to invoke a mechanism of identification to account for the complexity of what actually happens in practice? Can the psychoanalytic and behaviourist approaches be reconciled?

Perhaps because both views have been taken very seriously by child psychologists in the States, this dialogue has provided the impetus to a large number of ingenious and interesting experimental studies concerning the conditions under which young children feel anxious when tempted, avoid temptation, feel guilty after transgression, and model their behaviour on that of the adults around them. Many of these experiments are intriguing, and some, especially those which have produced unexpected results, have made a positive contribution to the understanding of child behaviour. This is a useful review because the literature in this field has now snowballed to such an extent that child psychologists can neither keep abreast of, nor afford to ignore, these studies. There is, however, a very real danger for students if knowledge of the research literature takes the place of more fundamental knowledge about how young children think and act, since the latter is essential in any attempt to interpret the results of these experiments critically. Unfortunately the findings are hardly ever clear-cut in the sense of being likely to force a radical revision of the theoretical constructs on which they are supposedly based; and perhaps this is because, in work of this kind, a large number of assumptions separate the initial theoretical postulates and what it is practicable to measure. As Mussen finally concludes, 'The nature of the problems involved is such that we can predict only gradual and continuous accretions to knowledge and understanding, since, for many reasons, the study of socialization does not lend itself to breakthroughs.'

'Risk Taking as a Function of the Situation, the Person and the Group', by NATHAN KOGAN and MICHAEL A. WALLACH

This last contribution is the most extensive of the three, and in fact occupies more than half of the book. It provides a systematic review of a large number of studies concerned with how people take risks in gambling situations both inside and outside the laboratory. In many ways, behaviour when gambling provides an ideal subject for the experimental psychologist: odds can be manipulated and controlled with great precision, and behaviour may be recorded very simply in terms of bets actually placed. There is, however, a fundamental difficulty in that the subjective feelings associated with winning and losing cannot be measured simply in terms of amounts of money won or lost, and the evidence points to the fact that it is these subjective motives which control actual behaviour. Whether or not a person will take risks depends upon: (1) how he evaluates fear of failure as against the sweet smell of success; (2) whether the subject regards the situation as a game or as 'serious'; (3) whether the subject believes skill is involved, and how skilled he in fact is; (4) whether or not he is showing off in front of other people; (5) what the

subject thinks the experimenter expects of him; (6) whether decisions are made by an individual or by a group.

What starts out as a hopeful new experimental area, with the possibility of highly controlled situations and mathematically precise predictions, thus slowly degenerates into a morass of doubt and speculation, the investigators balancing various complex human motives against one another in an attempt to retain just one or two valid generalizations which might be applicable to most persons in most risk-taking situations.

All three contributions are presented in a clear and interesting style, and each provides both a comprehensive account of a wide variety of empirical studies and a useful evaluation of them in the context of prevailing theoretical ideas. From this point of view, the whole book admirably achieves what it set out to accomplish. It raises, however, a fundamental issue: how far is the conventional 'hypothesis testing' research strategy successful in providing new and important knowledge which genuinely contributes to our understanding of human behaviour? What seems almost invariably to happen is that abstract theorizing (often quasi-mathematical) only leads to relatively 'clean' experiments, which confirm or invalidate specific hypotheses, if it is based upon a number of untestable assumptions. So long as the hypotheses are confirmed, we learn nothing about the shakiness of the original theoretical assumptions. Once contradictory evidence from other studies appears, this prompts a search among the theoretical assumptions to see whether they are, after all, really plausible. As time goes on, the theoretical assumptions are undermined one by one, and finally research workers are forced to recognize that in this situation, as in most others, real human behaviour is just about as complicated and difficult to predict and understand as most people had always supposed it was. The intelligent non-psychologist can perhaps be forgiven if after reading this volume in the hope that new directions might lead to new destinations, he feels that he has been taken for a ride.

JOHN NEWSON

Explorations in Typing Psychotics. Edited by M. LORR. Oxford: Pergamon Press. 1966. Pp. 241. 60*s*.

This book, intended as a sequel to *Syndromes of Psychosis* (Lorr, Klett & McNair, 1963), is divided into three parts. Part I, 'Problems in typing', consists of an introduction (unsigned), which defines the three concepts of symptom, syndrome and type, a chapter on 'Typing and taxonomy' (Lorr) and a literature review on typing methods (Lorr). Part II, 'Interview-based types', consists of five separate papers: 'Acute psychotic types' (Lorr, Klett and McNair), which identifies nine male and nine female types from the scores of 374 men and 448 women on the Inpatient Multidimensional Psychiatric Scale; 'Reliability of the acute psychotic types' (Klett and McNair), which presents eleven tables of various kinds of interrater correlation coefficients on the I.M.P.S.; 'Validity of the acute psychotic types' (Klett and Lorr), which presents thirteen tables of the frequency distributions of the various types on a range of variables from age, ethnic origin and religious background to prescribed treatment and outcome; 'Chronic psychotic types' (Lorr), which reports two I.M.P.S. studies identifying together fourteen types, ten of which are equated with previously found acute ones and four of which are new; and 'Types based on repeated measurement' (Rice and Mattson), a longitudinal study which with two kinds of analysis identifies nine and seven types. Part III, 'Ward and community behaviour types' (Lorr), which, using another rating scale, identifies seven types, six of which are equated with ones previously found; and 'Behaviour patterns of schizophrenics in the community' (Katz, Lowery and Cole), identifying six types.

In view of the present disordered state of our knowledge with regard to psychiatric diagnostic categories in general, and psychotic ones in particular, all investigations such as those reported here, applying the principles of numerical taxonomy to the symptomatology of mental illness, are of the highest importance. This having been said, it should be pointed out that this book consists of an unintegrated collection of papers, some of which would more properly fall into the category of interim internal reports (not for publication). For example, the literature review on typing methods is not followed up, and the typing methods used appear more or less arbitrarily selected. There is no overall attempt to relate to one another the many types identified and idiographically described: even when types from different studies are equated this is on subjective grounds. There is a plethora of indigestible data tables, particularly in the chapter on validity, where no statistical tests are used. It is very properly therefore that the authors at various places describe the work as 'preliminary' and 'tentative'.

J. P. N. PHILLIPS

Acquisition of Skill. Edited by E. A. Bilodeau. New York and London: Academic Press. 1966. Pp. xiii + 539. 100*s*.

The book comprises ten original papers, eight of them with comments, presented at a conference on the 'Acquisition of Skill' in 1965. Each contribution is a carefully prepared statement, in some cases a summary survey of a field, and in others a report of new experimental work or an attempt to formulate a theoretical viewpoint. This leads to some unevenness since some chapters will be of interest only to specialists and others will be useful for the student. The subject-matter is not easy to define, as Irion points out in his introductory history of research on the acquisition of skill. However, work on verbal learning is emphasized in several chapters and some topics which are at the very heart of skill research are unaccountably omitted.

Irion's review deliberately omits work in the last two decades in engineering psychology and concentrates on topics such as the distribution of practice, warm-up and interference effects. Names such as Woodworth, Lashley, Bartlett and Craik do not appear in the bibliography; the only reference to Paul Fitts is his 1951 review. Two chapters, one by Poulton on 'Tracking Behaviour' and one by K. U. Smith on 'Cybernetic Theory and Analysis of Learning', deal specifically with continuous skills whilst most of the others deal with discrete tasks where it is possible to talk of specific responses to specific stimuli and so make some sense of S–R terminology. Poulton's review, together with Briggs's comments, comprise a description of tracking sufficiently lucid to be a suitable introduction for students. Smith describes further experiments on the effects of the transformations of visual and verbal feedback using CCTV, video-tape recorder and a computer, and gives an account of a quasi-neurological (neuronic) theory of learning.

The paper by Jones, with comments by Fleishman, represents the tradition of individual differences in skills. The changing pattern of intercorrelations at different stages in skill acquisition was probably first noted by Woodrow 30 years ago, but this line of investigation has had a revival in recent years. It is often said that correlational studies are useful at the hypothesis-forming stage, and this work is of some help in drawing attention to the point that at different stages of learning trainees are doing different things and that different mechanisms are most probably involved. At this stage one surely needs to look beyond the correlations and attempt to analyse what is in fact happening.

Six of the ten principals in this symposium are Iowa allumni, sharing a basically S–R approach to skills, an approach very different from that which might be taken at say, Michigan or Cambridge. Noble's articles on 'Selective Learning and the Selective Mathometer' is an attempt to improve empirical laws of selective learning in multiple-choice situations. Adams, with comments by Alluisi, discusses 'Attention', including in his review some account of anticipation and information-processing limitations and vigilance. Although this was not intended as a compendious review it is hardly adequate as a summary of recent findings and current theories, particularly of work on this side of the Atlantic. Adams still feels it necessary to apologize for the use of the term 'attention' and feels it is ultimately moribund. Most of Battig's paper on 'Facilitation and Interference' deals with verbal learning and only in the last few pages does he deal with implications for research on transfer in motor-skill. Although every study of skill acquisition is strictly speaking a study of transfer, it is interesting to note how the recent training literature has almost abandoned the concept of transfer as, in the field, generalized training is rapidly falling out of favour.

Ina Bilodeau gives us a workmanlike review of work on information feedback, or knowledge of results, in the acquisition of simple graded responses which is nicely supplemented by Bourne's comments on concept identification; whilst E. Bilodeau, using the same lever-pulling task and a simple verbal task reports his recent work on retention. It is a pity that no reference is made to Nuttin's challenging work. E. Bilodeau uses the measure of intertrial consistency as an index of learning and retention in studies involving an impressive 11,000 subject sessions. Whilst as a definitive measure correlation may have its own problems, as Bahrick points out, one cannot help feeling that Thorndike's conclusions from his celebrated line-drawing experiment might have been rather different had he taken Bilodeau's approach.

The symposium is concluded by Underwood making some comparisons between motor skills and verbal-learning research. This was clearly not an easy task. Verbal learning may have more to learn from motor skills research than vice versa.

It is difficult to review adequately the work of eighteen different authors, particularly when there are divergences in theoretical orientation amongst them and between many of the authors and the reviewer. With one or two exceptions the papers taken individually represent useful summaries or provocative individual contributions. Taken together as a set of papers intending to define a field, the book is a serious disappointment. One might not have cause to complain about the emphasis on recent verbal learning research if it were not the case that, as Underwood almost admits, this field has remarkably little to offer in the way of concepts and methods to the study of skills. By pushing engineering psychology into the background the editor has deprived his readers of major sources of ideas which have done so much in the past two decades to stimulate basic research. The book is, of course, about the *acquisition* of skill. It could be argued that approaches inspired by informational concepts appear to be more useful in dealing with skilled *performance* than with skill *acquisition*, hence the need to use S–R concepts and methods which have been developed in the study of learning. It is the reviewer's opinion that this need not be so, but it has to be admitted that the kind of theory of skill acquisition put forward recently by Fitts, for example, has not yet had an impact on other learning theorists. In case the title proves misleading to those actually engaged in training, it is necessary to point out that the book makes no reference to recent work on the acquisition of real-life complex skills.

To sum up, the book is to the reviewer a disappointment because, despite the number and distinction of its contributors, it does not provide a balanced picture of recent research and recent thinking in this important field.

JOHN ANNETT

Macromolecules and Behaviour. Edited by JOHN GAITO. Amsterdam: North-Holland. 1966. Pp. ix + 197. 58*s*.

A feeling has arisen in the last few years that the time may be ripe for an approach to molecular interpretations of mental function. One man, Holger Hydén, is responsible for much of this feeling; a wide audience has been influenced by his speculations and, perhaps more important, by the techniques his laboratory has developed for the study of brain tissue.

An enormous amount of work and discussion has been going on, and the whole subject has become highly controversial. *Macromolecules and Behaviour* does not resolve these controversies; though, if anything, it leaves one with a sense that they have been underplayed, partly because, unlike another recent symposium (*Molecular Basis of Some Aspects of Mental Activity*, edited by Otto Walaas, 1966), it does not include questions or discussion. The book is a collection of papers which were originally delivered at a conference at Kansas State University in 1964, but some of them were revised for publication.

There is relatively little reference to the famous cannibalism experiments, where it was claimed that the transfer of RNA extracts from trained flatworms to untrained ones had a facilitating effect on learning in the untrained animals. Most of the work reported concerns the analysis of the RNA content of various parts of the mammalian brain, and the attempted correlation of RNA content with nervous activity.

Several authors cover the same ground here. One might especially mention the contribution of Pevzner. He feels it established that nervous excitation, of various kinds, induces an acceleration of RNA turnover in the brain, synthetic processes usually predominating over processes of breakdown. But, as, as Pevzner says, these changes need not be primary; they could be a compensatory response of the neurone to some other, completely different, process. Other lines of evidence are described: principally the administration of RNA in the human diet, with pre- and post-administration testing of memory; the use of inhibitors of protein and RNA synthesis; and the use of agents, like ribonuclease, which destroy RNA.

In general, the authors treat matters of molecules well, but the book is not fully satisfying when it deals with neurophysiology and behaviour. There is an inadequate critique of the work of Hydén, which depends on the isolation of supposedly homogeneous preparations of neuroglia. Further, there is insufficient discussion of the behavioural controls which would be needed before one could make unequivocal connections between the act of learning and changes in brain chemistry. The difficulty of some kinds of relevant experiment are enormous. For example, one wonders with Pevzner whether the slow fractionations and isolations of biochemists will ever be capable of following the fleeting pattern of events in a neurone. A good idea for future experiment is contributed by James Bonner. He outlines procedures, well tried in molecular biology, which

should resolve at least one of the current questions: whether or not the RNA synthesized by cortical neurones following learning derives its base sequence from the animal's stock of gene-carrying DNA. If the memory trace is coded in the molecular structure of RNA, then learning will be accompanied by the appearance of RNA of new and specific base composition. And this RNA would be detectable by its willingness to form stable double helices (i.e. the usual *in vivo* structure) with single-stranded samples of the animal's DNA. Karl Pribram speculates with verve and originality; for him the richness and subtlety of subjective experience, far from making molecular explanations of mental function irrelevant, make them necessary. He also considers neural holograms, and altogether welcomes the unorthodox.

The book, in general, is more biochemical than behavioural, but there is much in it that is bound to interest psychologists. The case that changes in brain RNA may be primary in learning and memory continues to look exciting but not proven, and no one can say how long one may have to wait for more.

NICHOLAS EVANS
HANNAH STEINBERG

Conflict in Society. Edited by ANTHONY DE REUCK and JULIE KNIGHT. London: Churchill. 1966. Pp. xii + 467. 65*s*.

There were in all 23 members of this international, interdisciplinary CIBA Foundation Symposium held in the summer of 1965. What we have here are 14 formal papers interspersed with ten discussion sessions. The official aim of the investigation was to discuss conflict between social groups at a hierarchy of levels.

The batting order of the disciplines represented by formal papers was roughly anthropology, ethology, psychiatry, sociology (strongly represented), with a substantial tail of political science, law and international relations. Psychology as such was not represented. This was perhaps an advantage, to psychology as well as to the symposium. So often in symposia on aggression, conflict and the like, the psychologist in the team is put in after ethology, and is then in the essentially negative position of having either to refute the claim of ethology to be able to generalize about man or, in anticipation of later discussions, to explain why he is reluctant to generalize to the international sphere from the modest conclusions of small-group experiments. But this symposium was blessed with an ethologist much less disposed than some to make extravagant territorial claims (thus, it is he, Michael Chance, who knocks down the attempt of a political scientist to apply directly to man findings about the ritual resolution by animals of potential conflict by means of rank-order formation, on the ground that ritualized submission denies to ninety per cent of human society the full use of intelligence); and with international relations experts whose analyses of international conflict point incidentally to the areas in which psychologists can best contribute to such discussions—by studies of popular attitudes and images, of decision-making processes and the effects of information overload.

In all, the symposium makes good reading. Different readers will, of course, be stimulated by different contributions. Here and there sparks are given off in the discussion sessions: among the formal contributions, this reviewer found most exciting the discussion of J. Van Doorn, the Dutch sociologist, of the conflict that exists in formal organizations between the needs of the individual and the requirements of the organization, so that the system benefits from a high degree of 'functional autonomy of parts'; Anatol Rapoport's presentation of contrasting views of war, cataclysmic and strategic, and stress on the ultimate dangers of the position taken by professional strategists; and John Burton (International Relations) analysing international relations in a way that makes more use of role concepts than of notions about human nature, and of systems analysis than of power conceptions.

THELMA VENESS

The Physiological Basis of Habituation. By E. M. GLASER. London: Oxford University Press. 1966. Pp. 102. 15*s*.

This book gives an account of Glaser's own experiment on cardiovascular response decrement to thermal stimuli, though the borderline between habituation and acclimatization remains undefined. There are also clear descriptions of changes in less studied responses such as motion sickness, and sweating after munching hot chillies. But it is not a general book on habituation: 23 of the 55 references are to the author's own work; Harris gave 123 references on habituation in

the *Psychological Bulletin* (1943). There is a similar personal flavour about the text, the author's work being reported and then noted as confirming Pavlov and other researchers.

Glaser's views on habituation are very different from those concurrently held by psychologists, since for him habituation is not a form of learning: '...the understanding of habituation was not helped by the assumption, still common among psychologists, that habituation is part of learning or "negative learning"' (p. 2). One wonders whether a physiologist in the U.S.S.R. or U.S.A. could so disregard the contributions, aims and terminology of psychologists. This might not matter if the book were at least accurate as far as it goes. But it is not. In a chapter entitled 'Properties of habituation' it is suggested that the number of stimuli, rather than the inter-trial interval between the stimuli, determine the degree of habituation. No attempt is made to examine the large number of experiments indicating that short inter-trial intervals give quicker habituation, even though the author refers on the same page to experiments by Coombs (misspelt Coombes) which specifically make this point.

Glaser's work on the effects of cortical lesions is of interest, but here again there is no background and no link with current theories, no mention of Sokolov, for instance. The uninformed reader is left with the impression that habituation is cortical. Prosser and Hunter's demonstration of habituation in spinal animals, reported in the physiological literature, is nowhere referred to.

From these shaky foundations the dazzling explanatory power of habituation in the last chapter suddenly carries the author away. He accounts for the success of the suffragette movement, the disappearance of fashions in time of war; and offers practical advice on race prejudice and television violence. Extreme speculation is balanced only by extreme caution: 'Any community that accepts habituation to what is not in its interest is probably exposing itself to risks' (p. 87). Probably? Anyone interested in habituation will do far better to study Thompson and Spencer's paper in the *Psychological Review* (1966, **73**, 16–43). N. MROSOVSKY

Principles of Psychophysiology. By RICHARD A. STERNBACH. New York: Academic Press. 1966. Pp. vii + 297. 60*s*.

This book is a hybrid; it contains 158 pages of text and the remainder is devoted to selected readings. It can be said to go some way towards filling a gap which is left by other undergraduate texts in physiological psychology. The author goes to some length to explain the subject matter of the sub-discipline of 'psychophysiology', and attempts to distinguish it from that subject matter covered by the field of physiological psychology. The appearance of this book and, for instance, the recent appearance of the journal *Psychophysiology* perhaps represent the coming of age of a branch of psychology which up to now has tended to be somewhat neglected, or even not wholly respectable.

It is a pity that this book—the first in the field—does not give a better impression. To say that it is introductory is an overstatement. It certainly covers the salient topics which are particular to the subject, e.g. 'stimulus-response specificity' and 'the law of initial value', but does this in a way which is at an irritatingly low level, rather than at the depth at which comparable standard physiological psychology texts are pitched. Perhaps in 158 pages one should not expect more, but a fuller list of references would undoubtedly have made up for deficiencies in the text.

Nine examples of journal articles from the field are given, extending from Cannon (1928) to Ax (1964). The choice of papers gives a reasonable glance at some of the activities of the psychophysiologist, and it would be unreasonable to complain about the choice. In spite of reservations, this book must undoubtedly be welcomed as filling a place in the undergraduate reading-list. One cannot but hope that it is the precursor of something more substantial. P. H. VENABLES

Experimental Parapsychology. By K. RAMAKRISHNA RAO. Springfield, Illinois: Thomas. 1966. Pp. x + 255. $9.50.

If it were not wholly uncritical and if it did not ignore evidence or arguments likely to throw doubt on the reality of psychical phenomena, this would be an excellent book. Its unfortunate bias is evident, however, even when attitudes to ESP are being discussed. Thus Hebb is stated to have admitted 'that he did not require any further evidence of psi but that he was worried because it did not fit into his rational scheme of things'. In the salient part of the passage

referred to, Hebb, in fact, wrote, 'Personally I do not accept ESP for a moment, because it does not make sense....'

The case for ESP is provided by giving a fairly detailed discussion of three major experiments: the Pearce–Pratt series; the Pratt–Woodruff series, and the Soal–Goldney experiments. Although each of these experiments has been extensively criticized in recent years, the author manages to elude such criticism, largely by ignoring it. In the case of the Pratt–Woodruff experiment, in answer to the criticism that a similar result could have been produced by a trick on the part of one of the experimenters, he tells us that this is a 'disgraceful argument' having no 'merit as a scientific criticism'.

Anyone who is unaware of the low standards prevalent in the reports emanating from the Duke laboratory is likely to be misled by Dr Rao's account. Thus he states: 'In the Turner–Ownbey series, in which the agent and the percipient were separated by over 250 miles, an average of 10·1 hits per run was obtained.' Since the chance expectation was 5 hits per run the reader may be impressed. He may not, however, be aware of the fact that the average of 10·1 hits per run was wholly due to very high scores in the first three of the eight days during which tests were made. During each of these three days one young lady, the guesser, by mistake sent her list of 25 guesses to the other young lady, the experimenter, instead of to Dr Rhine as had been planned. During these first three days scores were phenomenally high, but during the following five days when the two ladies were asked to send their records independently to Dr Rhine, the first young lady's psychic powers completely disappeared. This unfortunate lapse in the experimental conditions was reported by Dr Rhine in his first account of the experiment published in 1934. Dr Rao does not appear to consider it necessary to acquaint his readers with these facts or to discuss any of the conclusions they might have provoked. Perhaps he would consider such arguments as disgraceful and having no merit as scientific criticism.

Psychologists who are not already acquainted with experimental work in parapsychology are likely to be puzzled and misled by this book.

C. E. M. HANSEL

The Transfer of Learning. By HENRY C. ELLIS. New York: Macmillan. 1965. Pp. vii + 200. 16*s*.

This paperback offers a reasonably general introduction to the literature of transfer, with an emphasis on its application to education. In accordance with the format of the 'Critical Issues in Psychology Series' the author's text precedes a selection of reprinted papers so that, as the foreword has it, the 'combination will afford instructor and student opportunity to read stimulating, even challenging, argument with primary sources at hand'. In the present case there are only 80 pages of text, followed by papers from Underwood, Osgood, Harlow, Russell and Storms, Duncan, Ellis and Muller, and Hudgins. The usefulness of the reprinted material seems diminished by the virtual absence of linking passages or editorial comment other than that provided in the earlier exposition.

The text itself is uneven in quality, ranging from the relatively full discussion of a possibly peripheral topic like stimulus predifferentiation to the astonishingly superficial treatment of bilateral transfer. The eight lines devoted to bilateral effects might have been better omitted and constitute perhaps further evidence of the author's discomfort in the face of motor learning, otherwise implied by statements like '...negative transfer is difficult to obtain in verbal learning... On the other hand, negative transfer is frequently seen in motor skill learning where new responses must be attached to old stimuli.' The sketchy treatment of motor skills may be justified by the educational context, but the same argument can hardly apply to the experiments on retroaction. These are omitted entirely although Osgood, whose notorious surface relies heavily upon retroaction data, is featured prominently. The words 'retroaction' and 'retroactive' do not even appear in the index.

Again, there tends to be a looseness of articulation between different parts of the text. For example, the relatively heavy introduction to transfer designs and formulas is not carried through into the later evaluation of experiments. A more serious defect at the other end of the material is that a number of points emerge for the first time only in the concluding suggestions for teachers and the theoretical review. Thus, recommendation 3 appears to deal with the value of positive and negative instances in concept formation, a completely new topic. There is a mildly desperate search for theoretical issues—'...unfortunately, some researchers in transfer have been content

to establish functional relationships...without great concern for theoretical interpretation of such relationships'—in which the subjects of transposition and mathematical models suddenly occur. Nevertheless, the book does convey a good deal of relevant information, in parts with great clarity, and is likely to be widely used by undergraduates and training college students.

D. H. HOLDING

Genetic and Environmental Factors in Human Ability. Edited by J. E. MEADE and A. S. PARKES. Edinburgh and London: Oliver and Boyd. 1966. Pp. xi+242. 57*s*. 6*d*.

It is fashionable to smile at the nature–nurture controversy as it was fought a generation ago. No one now thinks that human ability can be described as either predominantly genetic or environmental in origin, but although it is a truism to regard both as important and to stress their interaction in the individual, it is nevertheless far from easy to investigate the precise mode of such interaction. This book, which is the happy product of a Eugenics Society symposium in October 1965, comes nearer to a detailed discussion than many others and will repay careful study by a wide readership.

It begins with a challenging section on the 'new look' in intelligence testing, in which Prof. Vernon discusses, with his customary lucidity, the wide range of test performance of children in various cultures and also has some critical comments to make on the shaky psychometric foundations of 'creativity tests'. He is followed by Bernstein and Young presenting an advance instalment of the first socio-linguistic findings so far yielded by the computer. Further results are now even more eagerly awaited, though it is to be hoped that something more penetrating than non-parametric analyses will be used, and that notice will be taken of Vernon's warning in the preceding pages that the gap between potential capacity and actual developed ability is much more difficult to measure than is often realized. Prof. Himmelweit also reports new data on the relationship between secondary school organization, as reflected in streaming and other policies, and early leaving—a serious source of intellectual waste in our population; she rightly stresses the need to study deviant rather than modal groups, a complex of interacting variables rather than the single variable favoured by some researchers. Among the most useful contributions is that by Wiseman reporting further developments of his Manchester investigations which examine in minute detail (though only through the medium of the correlation coefficient) the matrix of interrelationships between school performance, type of school and home conditions. His findings suggest that adverse environmental factors, most of which can be unambiguously specified, have their greatest detrimental effects on young children and on those of above average ability. His evidence underlines the need for schools to forge a closer link with homes, as recommended in the Plowden report.

Current research on subnormality is reported by Walker, discussing new work on prenatal and perinatal complications; these are more common among lower socio-economic groups, and are partly preventible by social and medical measures. He calls for objective behavioural studies of normal as well as 'at risk' neonates: psychologists can obviously make a useful contribution here but only a very small number are actively engaged in neonatal research in this country. Prof. Dent and Dr Cowie write clearly on biochemical and chromosomal aspects of subnormality, while Kushlick presents a blueprint of a complete reorganization of services for the subnormal, based on his thorough epidemiological studies in the Wessex region. The book concludes with a series of elegant genetic studies; some of these in their model-building assume a greater degree of measurement accuracy for intelligence tests than would be allowed by most psychologists: it seems doubtful if a difference of ten I.Q. points can be incorporated into a model involving alleles and gene loci. Other papers discuss the relation between differential fertility and intelligence and one by Huntley reports psychological and dermatoglyphic findings not only from 320 twin pairs but also from their parents and siblings.

The quality of the contributions is exceptionally high, and it is a pleasure to welcome the proceedings of a conference which must have been informative but far from pedestrian. Although the chapters are disappointingly brief, and the book somewhat expensive considering its brevity, the contributions are concisely and lucidly presented and make a refreshing change from some of the more turgid conference reports which are all too prevalent.

P. J. MITTLER

Attention. By A. R. White. Oxford: Blackwell. 1964. Pp. 134. 18*s*.

This is a philosophical book, but that is not to say that it is of no interest to psychologists. It is a major preoccupation of philosophers to help psychologists to clarify their concepts, and psychologists are indebted to them.

This book is concerned with a 'family' of concepts of attention—and it is a very large family, especially when we include all the cousins several times removed. In a brief review no more can be done than to note a few of the distinctions of interest to psychologists which Professor White draws.

It is argued that attention concepts (including hope, fear, expectation and many others) are 'object demanding'. They are like belief concepts. And this must be agreed. Only William James has suggested the possibility of objectless beliefs—and this only as a philosophical joke.

The distinction between object-demanding and objectless mental states remains obscure even after a century of argument. It is understandably so. Those who draw this distinction have to make it sharper than it actually is. They ignore the difficult cases. When, for example, a man, for fun, sits up all night in an allegedly haunted house, what is he attending to? He is alert (alerted) and expectant, but what is he expecting? He can only say, 'Expecting something out of the ordinary—some visual apparition, queer noises, odd movements of things in the room, something that brushes against one's sleeve, or perhaps sulphurous smells—just anything out of the ordinary.' Or take the case of the dog, apparently asleep, which begins gently to wag its tail. This behaviour is explained when a few moments later its master comes in. The dog had heard the characteristic sounds of his master's footsteps. The dog was listening *for* his master's footsteps before he could listen *to* them. There are clearly different kinds of directedness, different ways of being object-demanding. In the analysis of attention concepts, philosophers and psychologists are perhaps less concerned with the nouns and verbs than with the prepositions: attending *to*, listening *for*, being interested *in* or curious *about*. Once we start drawing distinctions there is no end to the distinctions we can draw.

One topic in this book which should be of special interest to psychologists is the group of legal attention concepts such as committing an offence inadvertently, behaving recklessly, or driving without due care and attention. The analysis of these concepts presupposes the more general and theoretical distinction between 'attending as a spectator' and 'attending as an agent'. There can be no doubt about the importance of this distinction. It has been made before by philosophical psychologists, e.g. James Ward and G. F. Stout. White amplifies and corrects their accounts; but even so some obscurities remain. The difference is not one between two co-ordinate species of a genus since attending as an agent entails attending as a spectator. A motorist driving with care and attention is not, of course, attending to his attention, nor even to his driving as such, but he is attending spectatorwise to the cyclists he wants to overtake and to the traffic signals. But he is doing so with certain 'dispositional tendencies'—which psychologists call 'sets', 'states of readiness', abilities and tendencies (and among tendencies statistical tendencies and 'felt' tendencies). Most of these can be described in Ryle's ungainly phrase 'mongrel categoricals' (so described no doubt because it is an ungainly concept). But there is more to it than that. This is another case in which when we begin to draw distinctions there is no end to the distinctions we must draw.

Prof. White's book, compressed and condensed though it be, is an essential text for all who are interested in the field it covers. We await from him 'further concerning the same'. C. A. MACE

Psychology for Psychiatrists. By C. G. Costello. Oxford: Pergamon Press. 1966. Pp. x + 328. Paper 25*s*. Cloth 35*s*.

The title of this book indicates clearly those at whom it is aimed and the author's wish in the preface that it, 'like sherry, will whet the appetite' would seem to indicate that it is meant as an introduction for the young psychiatrist. No insult to psychiatrists is intended by the suggestion that he has probably overestimated what they can absorb at this stage. The material is presented in a very concentrated form and the author has not made comprehension easier by his style of writing, which consists of profuse interspersions of references and dates, occupying approxi-

mately 5 per cent of the text (and a further 44 pages of index). This style, which may be necessary for a review article or a reference handbook, makes heavy going in what is meant to be an introductory textbook.

However, the opinion that the author has misjudged his audience should not be taken as a condemnation of the book, which has many good points. The contents cover many subjects in which there is a growing interest in clinical psychology. For example, there are separate chapters devoted to arousal, stress, anxiety and behaviour therapy, as well as chapters describing and evaluating various methods of psychological assessment. For readers who already have a background of general psychology, these provide a useful summary of some of the recent work done in these fields. It can therefore be recommended for use in training schemes in clinical psychology. The style of presenting all the information in the form of experimental findings has the advantage for such courses of providing a ready up-to-date bibliography for seminar use.

Trainee clinical psychologists and experienced psychiatrists will also benefit from Dr Costello's fairly successful attempt to link experimental findings with clinical problems and to bridge the gap between 'subjects' and 'patients' which is often ignored in books written from an experimental standpoint. He has also tackled, though less successfully, topics like groups, role, and the self, which although respectable in experimental social psychology are often regarded with suspicion in the experimental approach to clinical psychology, despite their obvious relevance.

In summary, this book is rather difficult to read, but it is up to date and offers the reader with a reasonable grounding in psychology a fair summary of certain experimental work with applications in the clinical field. RALPH J. MCGUIRE

Foundation of the Theory of Prediction. By WILLIAM W. ROZEBOOM. Homewood, Illinois: Dorsey Press. 1966. Pp. viii + 628. $10.50.

Rozeboom has published a 600-page treatise on the foundations of the theory of prediction. He points out that the concepts themselves are easy enough to grasp if an honest effort is made to do so, but that the symbols that express them cannot be learned instantaneously any more than any other language can be mastered by a single reading of a dictionary. Consequently, the reader is advised to refer frequently to the glossary of symbols on pp. 607 ff. These pages will soon become thumb-worn by any serious reader, as the book abounds with equations written in a rather cumbersome notation. For the expert, this is perhaps an advantage as he can grasp the author's meaning readily, but it makes the book more suitable for lecturers than students, despite the fact that it is written in a chatty good-humoured style.

The author keeps to his purpose in emphasizing the theoretical rather than the practical development of the subjects. There is nothing about it of the mixed statistical text-cum-workbook produced at times by Vernon in this country and by Cattell in the States.

The factorial methods described are largely based on the extraction of centroids, a technique which is useful from a didactic point of view, but rapidly falling into desuetude in the computer age. It has even been described by Kaiser as indecent.

There is an interesting and valuable appendix giving matrix derivations which will be of value to the growing body of students who have a working knowledge of the more elementary matrix operations, such as products, transposes and inverses, but lack acquaintance with partial differentiation, Lagrange multipliers and characteristic roots and vectors. F. W. WARBURTON

Test Theory. By DAVID MAGNUSSON. Translated by Hunter Mabon. London: Addison-Wesley. 1967. Pp. x + 270. 64*s*.

This book aims to give a theoretically and statistically clear introduction to basic test theory in the measurement of individual differences; it also gives models in the field of test construction, counselling, selection and classification, and discusses the methodology of dependability. The author also aims at writing a readable book for those who lack extensive training in mathematics and statistics, but he is perhaps less successful in achieving the last objective, for, although the coverage is excellent and the technique unobjectionable, the statistical treatment is in the middle rather than the elementary range of difficulty, at a level similar perhaps to Guilford's well known textbooks on statistics and psychometric measurement.

The author does succeed in keeping the reader up to date in recent trends in the measurement of reliability, validity, differential prediction and selection. One wonders whether some of the minutiae provided in textbooks in such fields as factor analysis are justified in the computer age, although it is still necessary, presumably, for students to hack their way through old, inefficient and inaccurate methods in order to gain some insight into the analysis of the data. Similarly, much of item analysis is outdated now that factorial methods for analysing hundreds of variables are available. On the whole, the book is clear, systematic and just the teeniest bit stodgy. (Rather unfortunately for the tyro, the formula for the product–moment correlation on page 42 has no minus sign in the numerator. This may result in some new records.) However, British students taking courses at the master's degree level will find it otherwise both contemporary and useful.

F. W. WARBURTON

Adoption—Facts and Fallacies. By M. L. KELLMER PRINGLE. London: Longmans. 1967. Pp. 251. 21*s*.

This volume should serve as a model for the kind of ground-clearing operation that must precede any major enquiry in the social sciences. The author and her team are to be congratulated on the enterprise and industry which, in the space of 20 months, resulted in this comprehensive review of British, American and Canadian literature on adoption since 1948. It was conceived as a prelude to the national survey of adoptive families now under way, and includes no original work. Nor is any searching criticism attempted, since its avowed purpose was to chart what is commonly regarded as an unknown field. It must be studied with these self-imposed restrictions in mind.

How much is in fact known about adoption? A bibliography of 448 items (of which 37 refer to work in progress) suggests that the field has not been totally neglected. This is not to say that adoption workers have solved all the practical problems which beset them from day to day. Anyone requiring decisive answers to such questions as the degree of permissible delay in placing a child; the ideal age or length of marriage in the adopters; or the most desirable family structure (does it matter whether there are or might subsequently be children of the marriage?), will not find them here. The evidence on these and many other recurring questions is indeed conflicting, and it would surely be unrealistic to demand a series of detailed prescriptions.

In the three introductory chapters the author has wisely set limits to the kind of questions which research is capable of answering. For example, it seems quite inappropriate to ask what proportion of adoptions are successful in the absence of agreed criteria for assessing natural families. But she draws attention to some important issues which are still 'wide open for investigation': the influence of pre-adoptive events, motives for adoption, and the merits of adoption as compared with absorption into the biological family.

It so happens that the reviewer began to investigate motivation while this bibliography was being compiled, and he can count himself fortunate to have had some of his homework done for him. Equally it seems clear that many other workers in child care or child development will have occasion to acknowledge their debt. Dare one hope that Dr Kellmer Pringle's example will inspire researchers in other fields to perform a similar service?

MICHAEL HUMPHREY

The Influence of Culture on Visual Perception. By MARSHALL H. SEGALL, DONALD T. CAMPBELL and MELVILLE J. HERSKOVITS. Indianapolis: Bobbs-Merrill. 1966. Pp. xvii+268. $2.95.

The appearance of a preliminary report on this study four years ago stirred up an area that had remained quiescent for about half a century. One therefore looked forward to the publication of the full account, and the expectations aroused earlier are not disappointed. What set it all off, we are told, was a controversy between the late Melville Herskovits, who was one of the first American anthropologists to 'discover' Africa, and his distinguished psychologist colleague, Donald Campbell. Herskovits argued that cultural differences were probably powerful enough to influence modes of perception, whilst Campbell was sceptical. They decided to mount an empirical attack on the problem, in which Marshall Segall took a prominent part.

The psychologists' conversion to Herskovits' view seems to have taken place before they

actually went into the field, possibly on the basis of studying the literature, lucidly surveyed in the first few chapters of the present monograph. At any rate, they set out to elaborate a theoretical formulation relating particular types of culture and ecology to specific illusion-response tendencies, on the assumption that different kinds of perceptual inference habits would be generated by different environments. Three hypotheses were put forward: first the 'carpentered world', concerning the effect of rectangularity in the environment, which is expected to enhance the Sander parallelogram and Müller–Lyer illusions; second, the 'foreshortening of receding horizontals', whereby open-plain dwellers should experience the horizontal–vertical illusion more powerfully than those whose horizon is circumscribed by forests, mountains or skyscrapers; lastly an effect of learning to interpret drawings and photographs is anticipated but not fully specified.

Some of the collection of data was undertaken by Segall, but since the cost of financing fieldwork in a variety of cultures would have been prohibitive, most of the work was done by anthropologists who agreed to help. Whilst this in itself is an ingenious innovation and an example that ought to be followed, it did entail some difficult problems. The material and instructions had to be prepared with special care, and results were screened for inconsistencies by a Guttman-type analysis. Altogether nearly 1900 protocols, including 13 non-Western samples, were obtained in the original study, and further data required subsequently are presented towards the end of the monograph.

How far do the findings support the hypotheses? As far as the 'carpentered world' is concerned, the fit is on the whole good. The foreshortening hypothesis seems, in the words of the authors, 'not drastically out of line with the data' and this appears to the reviewer to be a more realistic assessment than the less qualified one contained later in the conclusion. The results cannot throw much light on the third hypothesis, since information about the samples was confined to cultural and ecological characteristics of groups of people and information about individual experience was lacking; this might be a promising lead to pursue.

Among the most satisfying aspects of this report is the methodological sophistication displayed. Having shown the results, the authors stand back and take a critical look at them, considering a variety of alternative interpretations as well as conflicting evidence. These are dealt with systematically, either by further manipulation of the data or by instituting some special control experiments. There are of course limits to such 'ex post' procedures, and it is possible that they may have overlooked some important factors. More radically, one might think up some entirely different hypotheses that could have produced similar findings; here is one (admittedly with tongue somewhat in cheek): 'Cultures differ in the extent to which a given set of presentations of an illusion result in a response increment or decrement.'

It is of course one of the virtues of this monograph that it stimulates the reader to think about the issues raised, including some fundamental ones that border on philosophy. For instance, how can we distinguish 'between total failure in communication and total difference in perception'? This could well become a burning problem if and when humans encounter 'aliens'. In sum, this is an excellent book, which marks a new beginning and opens up exciting prospects for further research.

GUSTAV JAHODA

Mechanisms of Animal Behaviour. By P. R. MARLER and W. J. HAMILTON. London: Wiley. 1966. Pp. xii + 771. 113*s*.

It might be thought that this is another book by zoologists for zoologists, strong in those ethological overtones of disrespect and even abuse of psychology which make hackles rise. And in one sense this would be correct: the two authors are zoologists, but the result of their collaborative efforts is neither limited in interest to those in their own discipline nor is it over-critical of others.

Instead they are faithful to their title and concentrate on the comparative physiology of behaviour—the substrates of behaviour in the shape of the morphological features and processes which subserve it. To this end we are led through a series of chapters which constitute an up-to-date review of one of a number of relevant areas. Rhythmicity, locomotion, response to sensory stimulation of various kinds—chemical, visual, auditory—orientation, navigation and embryological development are some of the topics covered in this way. No terminological concessions are made to the non-biologist and the emphasis throughout is on the careful examination of the experimental evidence. The theoretical position taken—when one is explicitly stated—is in-

formed with a fine caution and is invariably middle-of-the-road. Thus, the importance of both endogenous and exogenous sources of stimulation is always stressed, the interaction of genetic and environmental determinants is recognized as crucial, and the balance between description and manipulative experimentation is always carefully weighed.

The result is a book which may be approached in the first instance with some scepticism regarding its relevance to the psychologist but which rapidly repays study. While the emphasis is often on work with invertebrates, especially in the earlier chapters, the researches described range widely, including some from primates, not excepting man. A truly comparative approach is thus achieved in some cases where several phyla have been studied. For a broadening of one's knowledge about non-human behaviour mechanisms the harvest is indeed rich. The temporal coding of the flashes of fireflies—for all the world like a series of lighthouses—the communication systems of ants, which may be as complex as those of the dancing bees, and the ways of fisher bats are but a few of the cases in point. It is perhaps this quality of surprise which will endear it to those senior students for whom it is intended. It could, for example, be confidently recommended as supplementary reading for courses on comparative psychology, given that no reliance was placed on the coverage of psychology as such. For example, the small sections on learning and conditioning are so minimal as to amount to ignoring these topics completely, and the section on psychogenetics, especially selection, is itself so selective as to be misleading.

Well produced, with a wealth of illustrations, a preponderance of up-to-date references—a fairly comprehensive sampling of which showed that some 38 per cent date from the present decade, and 41 more from the previous—and a refreshing, though not absolute, freedom from misprints, the book presents a revealing and striking picture of the spate of contemporary work in comparative physiology.

P. L. BROADHURST

Practical Training for the Severely Handicapped Child. By MILAN MORGENSTERN, HELEN LOW-BEER, FRANZ MORGENSTERN. London: The Spastics Society and William Heinemann. 1966. Pp. 134. 21*s*.

The aim of this small, unpretentious essay, first published in German in 1937, is clearly stated: to give practical guidance to parents and teachers faced with the task of educating a handicapped child. A brief chapter discussing the therapeutic concepts underlying the method is followed by a description of a variety of techniques for training manual and language skills. A wealth of illustrations is provided in the form of case histories and pictures of the apparatus used, some of which can be easily reproduced, while some has been made available by Paul and Marjorie Abbatt. The goal to which the method is directed is the development to the full of the retarded child's reduced potentialities, however limited, and the writers warn against the dangers of specifying in advance, or on the basis of a formal assessment, what may be achieved once a therapeutic relationship has been established and training begun.

The authors' theoretical orientation has its roots in psychoanalysis and Gestalt theory, while their approach is pragmatic and humane. Their chief concern is to outline for the teacher certain essential attitudes towards the handling of handicapped children rather than to specify a particular programme of instruction. Although the theoretical concepts (and hence vocabulary) tend to be psychodynamic rather than those of learning theory, certain principles are clearly implicit: the importance of arousal, the use of operant conditioning methods and the desirability of positive rather than negative reinforcement are chief among these. Nothing in the book strikes one as unsound nor are there any claims that the specific techniques presented are necessarily better than others; perhaps most important, there is no attempt to inspire an overly optimistic attitude to what can be achieved, even granted careful adherence to the authors' principles. At the same time, the case-histories testify to very significant improvement among some of the children they have taught.

In the foreword Professor Tizard writes, 'This book is so rich in psychological insights, and the authors are so inventive in their techniques, that I believe it will prove an enduring contribution to the literature, worthy to take its place beside the classical studies of Itard, Seguin, Montessori and Decroly.' It might be added that some of these insights are by now receiving experimental support in laboratory investigations of subnormal learning processes. The book should prove of great value as a practical guide to parents and teachers, and can also be recommended as background reading to all psychologists concerned with problems of training the severely retarded.

ANN CLARKE

The Perception of Brightness and Darkness. By L. M. Hurvich and D. Jameson. Boston: Allyn and Bacon. 1966. Pp. ix + 141. $2.50.

This book is one of a series treating topics in experimental psychology in some depth. The authors give a comprehensive coverage to the psychophysics of brightness perception—absolute and difference thresholds, changes over time, contrast and constancy phenomena, all are analysed in detail and the various theoretical formulations examined and compared.

The series is intended primarily for undergraduates but this particular volume may fall between two stools in this country—it is probably too weighty to be used routinely in a perception course but has not sufficient detail to act as a reference book on the field. It may be most useful to students who are carrying out an experimental project within this area or to those specializing in experimental psychology for a first degree. Psychophysics is a topic which most psychologists thankfully drop after their examinations are passed, but for those who want to brush up their knowledge of Ricco's law and the Purkinje shift, this book will make a useful and relatively inexpensive addition to their shelves.

Apart from the inevitable dryness of the subject matter (which the authors attempt to leaven by analysing the problems and then describing the various experimental attacks which have been made on them), perhaps the greatest shortcoming of the book is that insufficient space was available to include an account of work on the physiology and biochemistry of the visual system. Underpinning the phenomena in this way gives them a logic and coherence which are otherwise difficult to achieve.

F. I. M. Craik

Interpersonal Perception. By R. D. Laing, H. Phillipson and A. R. Lee. London: Tavistock. 1966. Pp. 179. 35*s*.

One or two psychoanalysts continue to produce ideas, based on their clinical material, which are valuable because they come from a more intensive analysis of individuals than is common in more rigorous research. In this case the work of the authors in family therapy has led them to an intensive study of the relationships between pairs of people, analysed in terms of mutual perceptions.

The basic method is to ask each member of the pair 720 questions, e.g. 'How would she think you have answered the following: "She hates me"?' The questions refer to six areas of interpersonal behaviour, and are asked at three levels—how A sees B, how A thinks B sees A (meta-perspective), and how A thinks B thinks A sees B (meta-meta-perspective). By combining the scores of A and B at different levels it is possible to arrive at indices of agreement, understanding, feeling understood, together with more complex notions such as feeling misunderstood incorrectly, etc. This is not so much a test, as a method for studying in depth the mutual perceptions of two people.

A small piece of validating data about the test is presented: 12 disturbed and 10 non-disturbed marriages are compared: there was significantly more misunderstanding for the disturbed partners. It would be interesting to know whether the meta-meta-perspective is of much importance; indeed it is doubtful how many people could understand this notion and answer meaningfully.

Some case-studies are reported, one at length, of the complex ways in which two people can become confused about one another, and the impossible predicaments that can be generated. Some interesting, highly speculative, but possibly testable, ideas are put forward about how 'spirals' of misunderstanding may develop: the general idea is that distrust generates further degrees of the same, e.g. 'Jack may reason: "Look at all the things that Jill is doing to try to prove to me that she loves me. If she really loved me she would not have to be so obvious about it and try so hard. The fact that she is trying so hard proves she is pretending. She must be trying to cover up her feelings—she must be trying to cover up her true feelings. She probably loves Tom."'

This monograph makes an interesting contribution to a hitherto neglected topic—the analysis of established relationships between people. It proposes a research technique and some novel hypotheses. It is, however, curiously out of touch with other work on person perception and social interaction, and the authors would be understood more accurately if they presented their material in a more readable way.

Michael Argyle

Questionnaire Design and Attitude Measurement. By A. N. Oppenheim. London: Heinemann. 1966. Pp. 298. 35*s*.

The basic weakness of this book is that it tries to do too much. In under 250 pages of actual text it attempts—despite its title—to be an almost complete manual of social research, with the inevitable result that most aspects are skimped, some to the point of emasculation. Just to illustrate the crowdedness of the canvas, the opening chapter is devoted to 'Problems of Survey Design'; it runs for 20 pages, at most 6000 words. Chapter 2 is even more compressed, coping with 'Problems of Questionnaire Design'—problem formulation, piloting, interview *vs.* mail questionnaire, open *vs.* closed questions, field coding and information loss—all this in 22 pages.

Chapters 4 to 8 will be the most valuable to the practitioner, outlining various methods of obtaining information in an interview or self-completion form, though they are so brief as to appear, at first sight, merely a catalogue of techniques. The weakest chapter is the last, in which the analysis process gets an elementary and quite perfunctory treatment. Analytical processes in practice are poor now—what a pity if the book should confirm bad habits.

Despite the brevity of the treatment, the book is sound, balanced and, within limits, practical. And in any event the mere listing of problems is a salutory reminder that they exist. Most of the 'methods' chapters are profusely illustrated by examples of scales, check-lists, etc. In the 140 in. of text in Chapter 4, 52¼ is taken up with examples, whilst in Chapter 7 (Projective Methods) 83¼ in. out of just over 200 consists of examples. Such an extensive use of examples is admirable—for the reviewer the best thing about the book—even though some don't carry much conviction, having the appearance of Aunt Sallies specially built to be knocked down.

Each chapter ends with a 'Further Reading' list, useful but very mixed. Some are only suitable for specialists (e.g. Hansom, Hurwitz and Madow on *Sampling*, or Cox on *Experimental Design*), whilst some things one might have expected are omitted, e.g. Kish on *Survey Sampling* and Cochran's paper on 'Non-Experimental Designs'.

Three irrelevant appendices complete the book. One is an extensive code list for occupational status. Another is a 'statistics test' of four items (fail this, and you need a refresher, says the author). By all means insist that social researchers are aware of significance testing, statistical inference, etc., but let us fervently hope that the psychological tests advocated have more than this degree of validity and reliability. The third appendix gives a specially constructed nomogram for testing the significance of differences between percentages. It is quite unnecessary. Excellent nomograms for this purpose prepared by Rosenbaum will be found in *Applied Statistics*.

The book, says the author, '. . . could serve as a textbook for research students and for practitioners in market research.' Research students would be expected to have access to other texts in which the practice of social research is dealt with more fully, whilst the market researcher normally works in a multidisciplinary team and hence has ready access to specialist advice. Nevertheless, there is a great deal of useful stuff here. It won't supersede Sølltiz *et al.*, Festinger and Katz, Moser, Hyman, etc., though it could be used in conjunction with them and it adds to them in including some new developments such as Kelly's repertory grid technique. The definitive work on attitude measurement remains to be written whilst texts on social research will always, I suspect, have to be compendia, contributed by experienced specialists.

B. P. EMMETT

The Neural Basis of Behaviour. By L. S. Woodburne. Charles E. Merrill Books. 1967. Pp. vi + 378. 94*s*.

In his preface, Prof. Woodburne states that his aim is to fill 'the gap that exists between very elementary texts and the medical compendiums' by combining in one volume enough neuroanatomy and neural physiology to give students a substantial indication of the behavioural significance of neural function and dysfunction. Unfortunately the author approaches this task mainly from the standpoint of a traditional neuroanatomist, so that the book is overloaded with neuroanatomical detail and curiously lacking in its discussion of basic neural physiology, and of the fundamental methodological problems involved in relating behaviour to neural functioning. The neuroanatomical bias of the book explains the bizarre allocations of space received by different parts of the nervous system, e.g. the cranial nerves are discussed at greater length than either the limbic system, the reticular formation or the hypothalamus. A book of 378 pages on

the neural basis of behaviour which devotes a mere 8 pages to the reticular formation and which barely mentions sleep mechanisms is obviously ill-attuned to recent research.

The author is at his best when describing the basic organization of the nervous system, but directly he steps into other fields (e.g. in discussing the evolution of the brain, or techniques for studying the nervous system, or even the physiology of vision) the text becomes too superficial to be of value. The policy of including no direct references in the text is also unfortunate, as it prevents the reader from readily following up interesting or contentious points; each chapter does contain a list of sources and additional reading but these are often too general to be of immediate assistance.

The book is profusely illustrated but the diagrams are often ill-chosen and badly reproduced, and some are indecipherable. However, there are some original line drawings by Brenda Sutherland which are useful in amplifying some parts of the text.

The very high cost of the book also requires comment, since, although its subject matter is not exactly duplicated elsewhere by a single book, it would be advisable to invest in other books rather than in this disappointing volume.

ANDREW TREACHER

The Scientific Study of Abnormal Behavior. By J. INGLIS. Chicago: Aldine. 1966. Pp. xiv + 256. $8.95.

Despite its title, this book is more like an anthology than a textbook. The author is a proponent of the view that single-case studies are exceptionally rewarding—and unduly avoided. This attitude has penetrated his writing and the result is an unusual book in which he discusses seven aspects of abnormal psychology in terms of the efforts of a few selected psychologists.

The section on thought-disorder is devoted largely to the work of Payne, the section on psychotherapy deals almost exclusively with Truax's studies, and so on. Apart from the fact that this method has serious limitations, Inglis's selection of topics and investigators will not satisfy many people. On the other hand, if one simply approaches this book as an intelligent commentary on selected aspects of abnormal behaviour, it can be read with enjoyment and profit.

The section on memory disorders covers research work carried out by Inglis and his colleagues over the past ten years. It is extremely well presented and he clearly has a thorough grasp of the subject. Shapiro's contribution to single-case studies is also discussed in a clear and interesting manner—with deserved credit being given for the honesty and self-criticism which is so notable a part of Shapiro's work. The weakest chapters are those on sensory-deprivation and on the effects of psychotherapy. A few sensory-deprivation studies are quoted in detail but not analysed in depth, and indeed, this is impossible unless one also considers the large number of investigations which are deliberately omitted. Truax's extended studies of psychotherapy receive very gentle handling and this chapter adds little to our understanding or assessment of this work. The two chapters on reinforcement are dated and most of this material (plus a great deal more) is available in other texts.

Although it is of uneven quality, this book can be recommended for two reasons: Inglis argues convincingly for the use of more single-case investigations and his chapters on memory-disorder, brain-injury and thought-disorder are useful contributions to our understanding of these abnormalities.

S. RACHMAN

Cognitive Consistency. Edited by S. FELDMAN. New York: Academic Press. 1967. Pp. xiii + 311. 68*s*.

The nine essays, based on contributions to a symposium, which make up this book are by workers in the field of consistency theory and addressed mainly to other workers in this field. A fairly comprehensive knowledge is assumed and no attempt is made to review the main experiments or theoretical ideas for readers unfamiliar with them. The editor describes the book as an attempt to stand back from the data, assess results and consider basic issues. After a general theoretical review by William J. McGuire there are two chapters on the role of motivation in consistency theory, one by Jerome E. Singer and one by the editor. Two chapters debate from the varying points of view of Elliot Aronson on the one hand and Milton J. Rosenberg on

the other, interpretations of the results of research into the effects of varying amounts of reward. Martin Fishbein contributes a lucid and thoughtful analysis of the relation between belief, attitude and behaviour. Karl E. Weick discusses 'task acceptance dilemmas'. Conceptual and empirical problems are considered by Albert E. Pepitone. Apart from a description by Abelson and Kanouse of experiments on the acceptance of verbal generalisation the book is almost entirely discursive.

The various chapters analyse aspects of a putative theory intended to embrace those ideas which have appeared in such various guises as 'balance', 'congruity' or 'dissonance' theory. The authors freely acknowledge many possible pitfalls to their endeavour, but, while admitting the difficulties of theory construction in this area, none appear to question the value of the enterprise. There is no attempt at independent evaluation by authors not committed to 'consistency theory' in one form or another.

The flavour of much of the book can be tasted from chapter titles such as 'Motivational aspects of attitudinal elements and their place in cognitive interaction', and 'Some limits of dissonance: toward a differentiated view of counter-attitudinal performance'. Some fruits of the arguments can best be presented again in the authors' own words, taken in the following extracts from three different summaries: 'Attempting to make reasonable assumptions about the nature of inconsistency motivation may prove to be a convenient heuristic; it may also, of course, be a dead end or a blind alley'; 'This paper has been an essay into the consistent use of an atomic conception of psychological matter'(!); and 'Task acceptance dilemmas suggest that the processes of cognitive change are a good deal more complicated than investigators have realized in their studies'. Between the titles and the summaries there is often little to hold on to; some authors attempt to catch an occasional straw of behaviourism.

The book leaves an impression of a few intriguing ideas having more than a grain of truth which have been played upon and expanded to suit so many different purposes that the resulting structure has many of the qualities of a balloon. Like some other attempts at psychological theorising, it is amenable to pulls from many directions. The tough-minded will have no compunction about delivering the puff to blow it away. When so many of the authors clearly recognise the difficulties and shortcomings of their attempts to build a theory of cognitive consistency, committing their misgivings to print at such length seems unnecessarily masochistic.

MARIAN ANNETT

Frontiers on Physiological Psychology. Edited by ROGER W. RUSSELL. New York and London: Academic Press. 1966. Pp. xvi+261. 68 *s*.

This book contains seven essays on contemporary physiological psychology by distinguished American research workers, preceded by a commendably brief editorial Introduction. The topics covered are highly specialized and dealt with in considerable detail, largely because the emphasis is on exploring the frontiers, as construed at present, and not on giving '. . . detailed maps of areas that have already become familiar territories' (p. 3). The essays are entitled: 'Integrative Properties of Neural Tissue' (Theodore Holmes Bullock), 'Neuronal Mechanisms of the Cerebral Cortex' (Sidney Ochs), 'Neural Processing of Visual Information' (Russell L. De Valois), 'Visual Mechanisms beyond the Striate Cortex' (Mortimer Mishkin), 'Physiological Mechanisms in Attention (the late Raul Hernandez-Peon), 'Neural Processing during Learning' (E. Roy John), and 'Biochemical Substrates of Behaviour' (Roger W. Russell). Each essay is self-contained, has its own list of references (except Bullock's, which quotes no one), and emphasizes papers published after 1960: over a half of the cited works belong to the last seven years. There is an Author Index and a Subject Index. The latter is in four sections, 'Anatomical or Structural Property', 'Behavioural Property,' 'Biochemical Property' and 'Electrophysiological Property' corresponding, so the explanatory note tells us, '. . . to the four properties of living organisms'. This index is cross-referenced and will, no doubt, add to the usefulness of this well produced and set out book.

It is not possible to review, even briefly, the diverse content of the book, since there is, necessarily, no really common theme and such integration of facts and ideas as there is takes place within individual contributions rather than between them. It is worth commenting, however, that each essay is a review of a field in which the author is involved, except that by Hernandez-Peon, which deals mainly with the author's own contributions to the topic. Mishkin's article

brings out clearly the importance for visual functions of the prestriate cortex. Russell's contribution reviews very neatly the biochemical approach and the importance of such research for the understanding of both 'genetically precoded behavioral patterns' and 'new behavioral patterns' as well as memory. This essay and that by John show how two approaches to the same problem can be simultaneously both confusing and illuminating.

The book is suitable for advanced seminars and will obviously provide much scope for discussion. The choice of frontiers, although arbitrary, is itself a matter of considerable dispute and one wonders where the 'new horizons', 'frontiers' and 'recent advances' will be in ten years time.

G. S. TUNE

The British Journal of Psychology

The British Journal of Psychology (Monograph Supplements)

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology (about 512 crown quarto pages per volume), *The British Journal of Medical Psychology* (about 350 crown quarto pages per volume), and the *British Journal of Social and Clinical Psychology* (about 320 crown quarto pages per volume) appear in quarterly parts, the subscription price of each, payable in advance, being £4. 10*s*. per volume post-free ($16.50 in the U.S.A.). Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to the Cambridge University Press, Bentley House, P.O. Box 92, 200 Euston Road, London, N.W. 1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 288 pages. The subscription rate, payable in advance, is £3. 10*s*. per volume post-free ($13 in the U.S.A.). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C. 4; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36 Essex Street, Strand, London, W.C. 2.

A series of *Monograph Supplements* is issued in connexion with *The British Journal of Psychology* but is not included in the subscription for that journal. Orders may be sent to any bookseller, or the Cambridge University Press, P.O. Box 92, 200 Euston Road, London, N.W. 1, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor W. Sluckin, Department of Psychology, The University, Leicester;

for *The British Journal of Medical Psychology* should be sent to

Dr T. Freeman, Dundee Royal Liff Hospital, By Dundee;

for *The British Journal of Social and Clinical Psychology* should be sent to

Miss Thelma Veness (Social Psychology Editor), Department of Psychology, Birkbeck College, Malet Street, London, W.C. 1, *or* to

Dr R. R. Hetherington (Clinical Psychology Editor), 6 Abercromby Square, Liverpool 7;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor R. J. Audley, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen;

for *The Monograph Supplements* should be sent to

Professor I. M. L. Hunter, Department of Psychology, The University, Keele, Staffs., whose decision regarding the publication of all Monograph Supplements will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1 September of the preceding year.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £6. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1 annually), *The British Journal of Medical Psychology* (£1 annually), *The British Journal of Social and Clinical Psychology* (£1 annually), *The British Journal of Mathematical and Statistical Psychology* (£1. 10*s*. annually), and *The British Journal of Educational Psychology* (£1 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, Tavistock House South, Tavistock Square, London, W.C. 1.

-3 APR 1968

The British Journal of Psychology *Volume 58, Parts 1 and 2, 1967*

CONTENTS

SIMON, HERBERT A.
An information-processing explanation of some perceptual phenomena *page* 1

DAY, R. H., BURNS, R., SINGER, G., HOLMES, V. and LETCHER, D.
Sensory spatial after-effects in relation to chronological age, mental retardation and schizophrenia 13

WOBER, MALLORY
Adapting Witkin's field independence theory to accommodate new information from Africa 29

GAULD, ALAN and STEPHENSON, GEOFFREY M.
Some experiments relating to Bartlett's theory of remembering 39

MCNULTY, JOHN A. and ISNOR, CATHERINE J. E.
Affective connotation of words as a factor in free recall learning 51

HOWE, M. J. A.
Digram-frequency and immediate memory 57

DALRYMPLE-ALFORD, E. C.
Repetition and immediate memory 63

WEINER, BERNARD and FELDMAN, PAUL
Information processing related to stimulus novelty and complexity in a signal detection paradigm 69

HAYWOOD, H. CARL and WACHS, THEODORE D.
Effects of arousing stimulation upon novelty preference in rats 77

HOLDING, D. H. and PRITCHATT, D.
Deutsch's model applied to a conditional maze 85

CORTEEN, R. S.
Basal conductance level and motor performance 93

DENNER, BRUCE and CASHDAN, SHELDON
Sensory processing and the recognition of forms in nursery-school children 101

KOCH, JAROSLAV
Conditioned orienting reactions in two-month-old infants 105

PAYNE, R. W., NETLEY, C. T. and SLOANE, R. B.
Rigidity, drive and conditioning in neurotics 111

FRITH, C. D.
The interaction of noise and personality with critical flicker fusion performance 127

PAYNE, ELSIE
Musical taste and personality 133

HASLAM, DIANA R.
Individual differences in pain threshold and level of arousal 139

LEVELT, W. J. M.
Note on the distribution of dominance times in binocular rivalry 143

BERGLUND, GÖSTA W.
A note on intelligence and season of birth 147

BURT, CYRIL
The genetic determination of intelligence: a reply 153

BROWN, ROBERT
Indiscriminate operationism: a reply to Keehn on 'Avoidance responses as discriminated operants' 163

PUBLICATIONS REVIEWED 167

Printed in Great Britain at the University Printing House, Cambridge

Volume 58, Parts 3 and 4 *November 1967*

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
ARTHUR SUMMERFIELD

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W. 1
AMERICAN BRANCH:
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

60s. net $11.00

Assistant Editors

JAMES HOGG PETER MITTLER

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements:

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this. Supplementary material of special interest may be lodged with the Technical Information and Documentary Unit, Department of Scientific and Industrial Research.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references, and all the pages numbered. Footnotes should be avoided. Articles should where possible be sub-divided with titled headings, following the pattern of articles published from Vol. **56** (1965) onwards.
A brief summary, suitable for use as an abstract, should precede the article.

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers' errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation; and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the authors' names and corresponding to an alphabetical list at the end of the paper. This list must give authors' names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.
For other general guidance, authors are referred to *The Journals of the British Psychological Society. Suggestions to Authors.* Cambridge University Press. Price 5*s*.
The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

LIVINGSTONE MEDICAL BOOKS

GUIDE TO PSYCHIATRY
MYRE SIM
Second Edition, 1,070 pages, 1,343 references, 60s.

THE ROOTS OF THE EGO
C. FRANKENSTEIN
298 pages, 72s. 6d.

RESEARCH IN INFANT BEHAVIOR: A CROSS-INDEXED BIBLIOGRAPHY
Y. BRACKBILL
294 pages, 65s.

METHODS IN PSYCHOPHYSIOLOGY
C. C. BROWN
376 pages, 132 illustrations, £7. 15s.

UNFINISHED TASKS IN THE BEHAVIORAL SCIENCES
A. ABRAMS, H. H. GARNER and J. E. P. TOMAN
282 pages, 60 illustrations, £5. 5s.

THE ACTUARIAL DESCRIPTION OF ABNORMAL PERSONALITY
P. A. MARKS and W. SEEMAN
340 pages, 70s.

CHILDBEARING. ITS SOCIAL AND PSYCHOLOGICAL FACTORS
S. A. RICHARDSON and A. F. GUTTMACHER
352 pages, 3 illustrations, 77s. 6d.

THE NATURAL DEVELOPMENT OF THE CHILD
A. H. BOWLEY
Fourth Edition, Fourth Reprint, 224 pages, 84 illustrations, 12s. 6d.

PSYCHOLOGY IN RELATION TO MEDICINE
R. M. MOWBRAY and T. F. RODGER, C.B.E.
Second Edition, 420 pages, 50 illustrations, 36s.

SLEEP AND ALTERED STATES OF CONSCIOUSNESS
A.R.N.M.D. Vol. 45
S. S. KETY
570 pages, 213 illustrations, £11. 15s.

DEVIANT CHILDREN GROWN UP
L. N. ROBINS
360 pages, 115 tables, Cloth £5. 7s. 6d. Paper 77s. 6d.

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD: NORMAL AND ABNORMAL
R. S. ILLINGWORTH
Third Edition Reprint, 396 pages, 160 illustrations, 37s. 6d.

LESSONS FROM CHILDHOOD
R. S. ILLINGWORTH and C. M. ILLINGWORTH
398 pages, 37s. 6d.

THE AFRICAN WITCH
M. GELFAND
244 pages, 25 illustrations, 30s.

THE MEASUREMENT AND APPRAISAL OF ADULT INTELLIGENCE
D. WECHSLER
Fourth Edition Reprint, 300 pages, 16 illustrations, 47s. 6d.

PSYCHOLOGY AS APPLIED TO NURSING
A. McGHIE
Fourth Edition, 356 pages, 11 illustrations, 25s.

E. & S. LIVINGSTONE LTD., 15-16-17 TEVIOT PLACE, EDINBURGH

BULLETIN DU C.E.R.P.

Bulletin d'Études et Recherches Psychologiques

(Trimestriel)

Sommaire du numéro 1 (janvier–mars) 1967, 16e année

RECHERCHES ET ÉTUDES PSYCHOLOGIQUES:

X. CUNY. La circulation de l'information dans un système élémentaire d'un service de transport.

H. DURUP. Recherches de plans d'expériences temporels à transitions exhaustives simples ou multiples.

J. LEPLAT. Incertitude temporelle et tâche ajoutée.

M. OLIVIER. Contribution à une introduction aux études des systèmes hommes-machine.

ANALYSES DE LIVRES.

Sommaire du numéro 2 (avril–juin) 1967, 16e année

FRANÇOISE LANTIER et NICOLE MANDON. L'évolution des structures professionnelles dans la métallurgie lyonnaise: incidence des pratiques d'embauche et de formation.

Rédaction et administration: 13 RUE PAUL-CHAUTARD, PARIS (15e) SUF. 59–77

Abonnement (un an): France, 15 fr.; Étranger, 20 fr.
Le numéro: France, 4.50 fr.; Étranger, 6 fr.

Règlements au nom de:
A.F.P.A., CENTRE D'ETUDES ET RECHERCHES PSYCHOTECHNIQUES
13 RUE PAUL-CHAUTARD, PARIS (15e)
C.C.P. PARIS 9.170.53.

ZEITSCHRIFT FÜR EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

Herausgeber für Deutschland

Prof. Dr. Ph. Lersch, München; Prof. Dr. F. Sander, Bonn
Prof. Dr. K. Gottschaldt, Göttingen

Herausgeber für Österreich
Prof. Dr. H. Rohracher, Wien

Herausgeber für die Schweiz
Prof. Dr. R. Meili, Bern

Redaktion, Dr. C. J. Hogrefe, Göttingen

Band XIV-2

Festschrift für Theodor Scharmann

Georg Dietrich. Durch welche erzieherischen Maßnahmen kann die Einstellung von Volksschülern gegenüber 'dem Juden' verändert werden?

Fred E. Fiedler. Führungsstil und Leistung koagierender Gruppen.

Rainer Fuchs. Motivation und neurophysiologische Steuerung des Humanverhaltens.

Dieter Hanhart. Strukturelle Aspekte im dyadischen Rollensystem.

Helmut Klages. Über die Verwendbarkeit von Modellen in der soziologischen Forschung.

Ferdinand Merz. Über den Einfluß von Bekenntnisschulen auf die Werthaltungen Jugendlicher.

Heinz Alfred Müller. Aggressionen im Betrieb und ihre Bewältigung.

Hans Thomae und Heribert Simons. Formen der Reaktion auf Belastungssituationen. Ein Beitrag zur pyschologischen Alternsforschung.

Walter Toman. Motivation als Systemgeschehen.

Wilhelm Witte. Einige Probleme beruflicher Entwicklung sub specie der psychologischen Grundlagenforschung.

Band XIV-3

Willibald H. Butollo und Joachim Maly. Anpassung an die verzögerte Rückmeldung der Lautsprache in einem 3-Tage-Versuch.

Wassilios Fthenakis. Die Urteilsbildung in kleinen Gruppen unter Berücksichtigung des entwicklungs- und völkerpsychologischen Aspektes.

Günther Guttmann. Periodizitäten im Frequenzverlauf der hirnelektrischen Ruheaktivität.

Klaus Holzkamp und Peter Keller. Seriale und dimensionale Bedingungen des Lernens der Größenakzentuierung: Eine experimentelle Studie zur sozialen Wahrnehmung.

Günter Lehmann. Die Analyse geometrisch-optischer Täuschungen durch Vektorfelder.

Gerald M. Murch. Die Nachwirkung unterschwelliger Reizung als eine Funktion der Zeit zwischen Reizdarbietung und Reaktion.

Richard Olechowski. Experimente über den Stimm- und Sprechausdruck beim Lügen.

Leo Reiffenstein. Experimente über die Konditionierbarkeit der Mikrovibration.

Reinhold Sielle. Über die visuelle Berufsorientierung Jugendlicher.

Anne-Marie Tausch, Reinhard Tausch und Bernd Fittkau. Merkmalszusammenhänge der verbalen Interaktion und kritische Überprüfung typologischer Verhaltenskonzepte.

Günther Guttmann und Josef Gangberger. Subliminale Stimulation im menschlichen Thalamus zur Auslösung von konditionierten verbalen Reaktionen.

All articles are provided with summaries in English and French.

The journal appears quarterly and contains 160 pages per number.

Subscription is DM 64.— per volume.

VERLAG FÜR PSYCHOLOGIE . DR. C. J. HOGREFE . GÖTTINGEN
(Brentanoweg 10, Germany)

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA

YEAR XXVIII. ISSUE I — JANUARY–FEBRUARY 1967

CONTENTS

LEADING ARTICLE

PAPERS

L. ANCONA. The Institute of Psychology at the 'Università Cattolica del Sacro Cuore' (from 1959 up to 1966).

M. KLEIN. Our adult world and its roots in childhood.

YEAR XXVIII. ISSUE III–IV — MAY–AUGUST 1967

CONTENTS

Proceedings of the VIIIth Information Course on psychotherapy problems (Arenzano, 31st October–5th November 1966)

LEADING ARTICLE

PAPERS

L. ANCONA. Evolution of techniques for psychotherapic approach to schizophrenia.

G. BENEDETTI. Communication in psychiatry.

S. ARIETI. A rapid survey of studies about schizophrenic thought from Bleuler up to this time.

S. ARIETI. Secondary process in psychodynamics of schizophrenia.

R. SPIEGEL. Primary process: language and individual experiences in schizophrenic individuals' life.

S. SHEINER. Troubles of thought process in schizophrenia.

G. CHRZANOWSKI. Dynamics and genetics of thought troubles in schizopathic individuals.

M. SELVINI PALAZZOLI. Thought disorders and family relations in schizophrenic individuals.

S. PIRO. Language and communication in psychopathology.

L. FRIGHI. Language and a communication in psychotherapic dealings.

Every article is followed by an Italian, French, English and German summary.
Every issue summarizes many Italian and Foreign works of psychology, neurology and psychiatrics.
Annual subscription: Italy, 3,000 Lire; foreign countries, 6,000 Lire (or equivalent foreign money).
Redaction — Administration: 1, Largo Gemelli, *Milan 20123* (Italy).

A System of Computer Programs for use in Psychology and Education

by

H. J. HALLWORTH

assisted by

ANN BREBNER

This handbook explains the preparation of data for a system of computer programs which will perform a variety of statistical computations frequently used in the social sciences.

The separate programs were originally developed in the Education Department at the University of Birmingham, and are published by the British Psychological Society on the recommendation of a committee appointed to consider the development of computer programming. The programs have been organised into a system which may be used by those with little or no previous experience of electronic computers. It is merely assumed that the user is acquainted with a teleprinter or a card punching machine.

Price 10*s*. (post free)

Available from—

THE BRITISH PSYCHOLOGICAL SOCIETY

TAVISTOCK HOUSE SOUTH, TAVISTOCK SQUARE, LONDON, W.C.1

SUGGESTIONS TO AUTHORS

A NEW PUBLICATION OF THE BRITISH PSYCHOLOGICAL SOCIETY

This booklet contains advice on the preparation of papers for submission to journals published by the Society, and gives general information as well as notes on the arrangement of articles; preparation of manuscripts; mathematical and statistical formulae; recommended abbreviations and symbols; recommended spelling of commonly used terms; proof correction; a list of periodicals and their abbreviations and notes on copyright arrangements.

Available from

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W.1
32 EAST 57TH STREET, NEW YORK, N.Y.10022

5*s*. per copy (U.S.A. $1.00) post free

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 197–203
Printed in Great Britain

SERIAL AND PARALLEL CLASSIFICATION

By D. W. J. CORCORAN

Medical Research Council, Applied Psychology Research Unit, Cambridge, England

The recognition of complex stimuli may be achieved by either of two strategies. Stimulus characteristics may be noted in turn (*serial* classification) or simultaneously (*parallel* classification). Two experiments are reported which attempt to determine which method is adopted with novel stimuli and whether a change in strategy occurs with practice. It was concluded that an initial mode of classification which was probably serial gives way to parallel classification after extended practice. The advantages of parallel classification in recognition are discussed.

In discussing the problems involved in pattern recognition by computer, Selfridge & Neisser (1960) distinguish two modes of stimulus classification. In *serial* classification each distinguishing characteristic of the stimulus is treated in turn whereas in *parallel* classification the characteristics are treated simultaneously. Selfridge and Neisser deal specifically with the recognition of written letters and discuss how specific letter features may be dealt with in turn or simultaneously in order to arrive at a decision about the stimulus as a whole. Neisser (1963*a*, *b*), concerned with the serial *v.* parallel distinction in human recognition, had subjects search among lists of letters for one, or one of two different letters and found that searching for one of two letters was as efficient as searching for one alone. On the basis of these and other similar findings he concluded that a parallel strategy had been adopted. Neisser's situation was not, however, strictly comparable to the task presented to the computer in at least two respects. First, no attempt was made to analyse the process of recognition of a *single* letter, instead it was demonstrated that *two* processes of recognition (which may have been serial) could go on simultaneously. Secondly, letter recognition is greatly over-learned by human subjects; novel stimuli may well be classified serially and a change in strategy may occur when the stimuli become familiar.

The aims of the present investigation were twofold: (1) to determine which strategy is adopted with novel stimuli, and (2) to discover whether a change in strategy occurs when subjects become better acquainted with the stimuli.

If two stimulus parameters A and B are considered, where both parmeters can have one of two values, $A+$ or $A-$ and $B+$ or $B-$, a population of four stimuli, $A+B+$, $A+B-$, $A-B+$, and $A-B-$, may be constructed. When a subject responds appropriately to one of these four stimuli, has he considered A before B (or B before A) or has he considered A and B simultaneously?

In Fig. 1 (S) two discrimination mechanisms, a and b, are joined in series. Sensory input AB first enters a, is discriminated and the output A' (a correct or incorrect response by a) enters a decision mechanism D, where it is temporarily stored. The sensory input AB then enters b, and the output B' joins A' in D. The times taken for the sensory inputs to be discriminated by a and b are t_a and t_b respectively; thus the interval between presentation of the stimulus and the arrival of the information at D takes (t_a+t_b).

Fig. 1 (P) shows a and b connected in parallel. D will receive both inputs either in

time t_a or t_b, whichever is the longer. D will therefore have received sufficient information to respond to the total stimulus more quickly if the P system rather than the S system has been operating.

Suppose that D is instructed to respond in one way, X, if it receives information that stimulus $A+B+$ was presented, but in another way, Y, if stimuli $A+B-$, $A-B+$ or $A-B-$ were presented. Consider first the S arrangement. When presented with $A+B+$, D will receive inputs from a and b at times t_a+t_b, and since it cannot respond before receiving A' and B' it must wait t_a+t_b before being in a position to respond X. If $A-B+$ is presented, however, D will receive sufficient information to respond Y after time t_a or after time t_a+t_b (if AB enters b before a). If in a group of subjects (or within a single subject) a is in some cases entered before b and in some

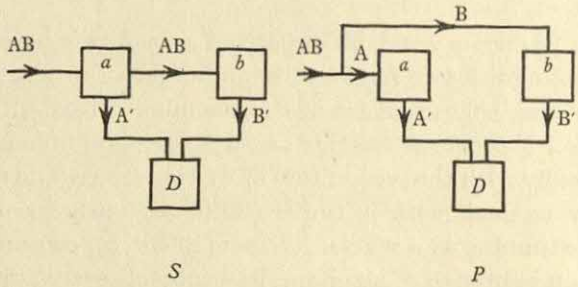


Fig. 1. Diagrammatic representations of a serial classifying system (S) and a parallel system (P).

cases after b, D will on average have to wait less time before it can respond, than if $A+B+$ had been presented. A similar reasoning applies to stimulus $A+B-$. When $A-B-$ is presented, D will have to wait either t_a or t_b, depending upon whether a is entered before or after b. Thus if a serial stategy is adopted, D will receive sufficient information to respond to $A+B+$ in time t_a+t_b, to $A+B-$ or $A+B+$ in a time less than t_a+t_b, whilst D will have to wait either t_a or t_b (a still shorter time) if $A-B-$ is presented. In the parallel system D will receive sufficient information to respond in either time t_a or t_b, whichever is the longer, no matter which of the four stimuli has been presented.

In Fig. 2 the time taken before D receives sufficient information is plotted for each of the four stimuli. Plot S1 in Fig. 2 shows the kind of curve which would be expected if a series system were operating and plot P1 if the system were parallel.

Two points may be raised at this stage. (1) If all the subjects in a group consider B before A at all times, D will receive sufficient information to respond to $A+B-$ in time t_b, which will also be the time required for stimulus $A-B-$. Thus a plot like S2 could occur as a result of a series classification. (2) If t_b were less than t_a and a parallel system were operating, D would receive B' before A' and thus have sufficient information to respond, when presented with one of the stimuli $A+B-$ or $A-B-$ in a shorter time (t_b) than if stimuli $A-B+$ or $A+B+$ were presented (t_a). In fact $A-B+$ or $A+B+$ would exceed $A+B-$ and $A-B-$ by (t_a-t_b). This plot is labelled P2 in Fig. 2. P2 resembles S2 in form and some difficulty in distinguishing the two may be expected.

Up to this stage only the time taken for information to *reach D*, has been considered, but what of *D* itself? Clearly before any predictions can be made about overt performance, the characteristics of *D* must be considered. *D*, which has been instructed to initiate response *Y* as soon as it receives any negative input, can reasonably be assumed to have one of two characteristics. It may be 'all or none', in that it will respond as quickly to one negative input as to two. This will be called a type 1 *D*-mechanism. If *D* is of type 1 the plots in Fig. 2 will conform to actual reaction times, since the time taken for information to *reach D* will be the only source of variability. If on the other hand *D* is of type 2 (which would cause it to respond faster to two negative inputs than to one), a plot like S1 will occur whether classification has been serial or parallel.

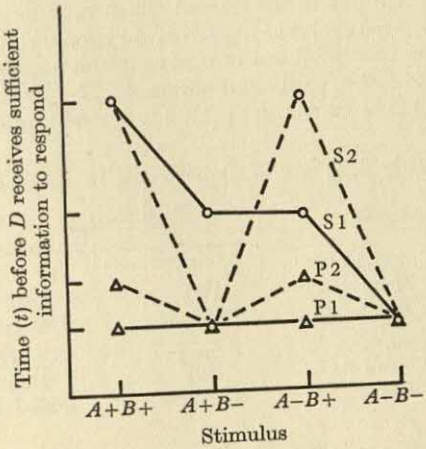


Fig. 2. Theoretical plots for serial strategies (S1, S2) and parallel strategies (P1, P2): S1, O——O; S2, O- - -O; P1, Δ——Δ; P2, Δ- - -Δ.

Table 1. *Type of curve to be expected under each of the four possible contingencies*

Mode of operation of *D*	Type of classification: Serial	Type of classification: Parallel
Type 1*D*	S1 but possibly S2	P1 or P2
Type 2*D*	S1	S1

Table 1 summarizes the conclusions reached up to this stage by showing the four possible contingencies. It will be noted that a P1 or P2 plot may only occur with a parallel system, whereas S1 or S2 may occur with either system, depending upon which type of *D*-mechanism has been served with the information. It is also clear that in order to distinguish the two types of classification a type 1 *D*-mechanism must be functioning. The first stage in the investigation therefore is to determine which type of *D* existed for the particular kind of stimuli to be used in the main experiment. (It seems likely if the 'plus' and 'minus' values on a pair of parameters are easily distinguishable that the subject will respond as quickly to one negative

value as to two. When the differences are less discriminable, however, subjects may react more quickly if they receive two pieces rather than one piece of negative evidence. The appropriate experiment to test these hypotheses is under way.)

EXPERIMENT I

Method

Stimulus material. The stimuli consisted of three packs of fourteen cards, measuring $3\frac{1}{4} \times 2\frac{1}{4}$ in. and containing a grey square on a black ground. The squares differed in area (the A dimension) and in brightness (the B dimension). $A+$ measured 1·6 sq. in. and $A-$ 1·4 sq. in.; $B+$ had a reflectance of 40 % and $B-$ a reflectance of 29·4 %. The stimuli were all photographs of a single grey square. The difference in area was achieved by printing at different sizes and the brightness was varied by doubling the exposure time. By using these techniques it was possible to prevent the occurrence of any small features in the stimuli which may have been used as cues in preference to their proper characteristics. Since negatives vary somewhat in density it was necessary to select the required number of stimuli out of a large initial population.

Pack 1 contained seven $A+B+$ cards and seven $A-B-$ cards; pack 2 comprised seven $A+B+$ cards and seven $A+B-$ cards, and pack 3 was made up of seven $A+B+$ and seven $A-B+$ cards.

Table 2. *Time (sec) per card in Expt I*

Subjects	Stimulus	
Group 1	$A-B-$	$A+B-$
Mean	1·11	1·07
Range	1·00–1·31	0·99–1·39
Group 2	$A-B-$	$A-B+$
Mean	1·22	1·22
Range	0·95–1·38	0·95–1·48

Subjects. Six naval ratings and six laboratory personnel served as subjects.

Procedure. The ratings (group 1) were required to sort packs 1 and 2, half receiving pack 1 before 2 and half having the reverse order. The laboratory subjects (group 2) sorted packs 1 and 3, with 1 presented before 3 to half the subjects and 3 before 1 to the other half. Subjects were required to place all $A+B+$ into an aperture measuring 4×2 in. and all other cards, whatever their type, into another identical aperture. An example of stimulus $A+B+$ was displayed near one aperture and either $A-B+$, $A+B-$ or $A-B-$ was placed near the other aperture, depending upon the particular condition.

The subjects were presented with a shuffled pack and informed that it comprised equal numbers of the two types of stimulus displayed on the apparatus. They were then requested to sort the cards with complete accuracy, but as rapidly as possible. The total time taken to sort the pack was recorded on a stop-watch. Each subject sorted both packs three times and the mean RT per card was computed.

RESULTS AND DISCUSSION

Table 2 shows the mean reaction time per card. It is clear that a pack containing $A-B-$ stimuli took about as long to sort as one containing $A-B+$ or $A+B-$. This result would apparently indicate that when D receives two 'bits' of negative information it reacts no faster than if it receives one, when stimuli of the present type are used. Hence it may be concluded that a type 1 D-mechanism was probably operating with the present stimuli.

Experiment II

Method

Stimulus material and apparatus. The stimuli used in Expt I were used in a single pack of sixty cards. The cards were sorted through apertures measuring 4 × 2 in. into a four-compartment box. As a card passed through an aperture it interrupted a beam of light and thus activated a photocell, which in turn caused an oscillator to react. A separate oscillator was activated by each photocell and each gave different outputs of 300, 500, 700 and 900 c/s. These tones were recorded on magnetic tape and later fed into SETAR (Welford, 1952). In this way both the response made and the reaction time to each card were recorded.

Subjects. Six naval ratings served as subjects.

Procedure. The working of the apparatus was first described briefly, then the subject was shown an example of each of the four cards and told how they were constructed from the two dichotomized dimensions of brightness and size.

The remaining instructions were as follows: 'This pack contains examples of each of the four cards. What you have to do is to pick out all the ones which look exactly like this [he was shown card $A+B+$] and sort them into this box which is marked with an X. All other cards you are to reject by sorting them into this box which is marked with a 'Y'. [The aperture was situated on the subject's immediate right and Y on his immediate left.] The first and last cards in the pack will be blank. The first blank card you should place down this hole [the aperture on his extreme left] and the last card down this one [the one at his extreme right]. The first card will send a low tone on to the tape when you start and when you finish the last card will send in a high tone. You should sort the pack as quickly and as accurately as possible.'

Thirty trials were conducted over five mornings and the five intervening afternoons. During each session the pack was sorted three times. The 3rd, 6th, 18th, 27th and 30th trials were analysed. A final trial (trial 31) was run immediately after trial 30. Subjects were requested to sort the cards independently of the stimulus they contained, placing three cards in Y and every fourth in X. These data were intended to give some indication of the time taken over the purely motor aspects of the task.

Results and discussion

Fig. 3 shows the mean reaction time to each of the four stimuli $A+B+$, $A+B-$, $A-B+$, and $A-B-$ on trials 3, 6, 18, 27 and 30. It will be recalled that only a quarter of the stimuli were placed in box X, so that there was likely to be a heavy response bias towards box Y. Since the X response is the least probable, longer reaction times to $A+B+$ may be expected irrespective of the validity of the foregoing argument. The reaction times to $A+B+$ must therefore be disregarded in all the plots of Fig. 3. The other three points, however, furnish sufficient information for several conclusions to be drawn.

Trials 3 *and* 6. The plots for trials 3 and 6 are similar in form to plot S1 of Fig. 2, but might conceivably be examples of one of the other types of curve. Plots S2, P1 and P2, however, have a single common feature not shown by S1: the reaction time to $A+B-$ is of the same magnitude as that to $A-B-$. In order to reject the hypothesis that plots 3 and 6 are examples of S2, P1 or P2 it is necessary to show that the reaction time to $A+B-$ is greater than that to $A-B-$. This was found to be the case for trial 6 ($P < 0{\cdot}05$); but the difference was not statistically significant for trial 3, although only a single subject showed the reverse trend. Since it was shown in the first experiment that D did not react faster to two negative 'bits' of information than to one, the evidence favours the conclusion that subjects were using a serial strategy in trials 3 and 6.

Trials 27 *and* 30. In trials 27 and 30 the reaction time to $A+B-$ is of the same

magnitude as that to $A-B-$, but the reaction time to $A-B+$ is slightly higher than that to the other stimuli. These patterns could at first glance be examples of plots S2, P2 or P1. In S2 and P2, however, the reaction time $A+B-$ is equal to that for $A-B-$ and lower than that to $A-B+$. The S2 plot occurred when all subjects at all times considered the B parameters before A; hence the reaction time to $A-B+$ must be reliably higher than the other two. P2 was the result of a parallel process in which the B parameter was discriminated faster than A ($t_b < t_a$), so that in this case also the reaction time to $A-B+$ should be reliably higher than that to the other two stimuli. Since the difference between the reaction times is small in the P2 case ($t_a - t_b$) one might not expect a significant difference and it is therefore difficult to distinguish P2 and P1.

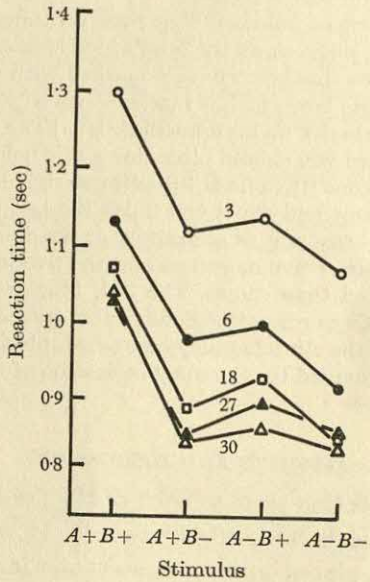


Fig. 3. Reaction times to the four different stimuli on different trials in Expt II: trial 3, ○——○; trial 6, ●——●; trial 18, □——□; trial 27, ▲——▲; and trial 30, △——△.

It is necessary to the argument, however, if possible to be able to reject the hypothesis that the data of trials 27 and 30 could be of the S2 type. The reaction time to $A-B+$ was found to be not reliably different from that of either of the other stimuli. (On trial 30, for example, the mean RT to $A+B-$ and $A-B-$ was lower than that to $A-B+$ for three subjects, was greater for two and identical for one.) Thus trials 27 and 30 are probably of the P1 or P2 type, and suggest that in these trials parallel strategy had been adopted. There is the possibility, however, that the apparatus imposed lower 'motor' limits on the reaction time, so that the 'real' shape of plots 27 and 30 was similar to that of 3 and 6. This cannot be the case, however, since on trial 31 in which the lower limits of the apparatus were determined the mean time for the X box was 0·52 sec and for Y 0·51 sec.

General discussion

The experiments have shown (*a*) that when subjects are dealing with novel stimuli they are probably using a serial strategy, (*b*) that with training a parallel mode of classification probably develops.

It seems plausible for at least two reasons that parallel processes are utilized in human recognition. First, parallel classification takes less time. A complex stimulus such as a face or a voice is immediately recognized when it is well learned. It seems unlikely that each distinguishing characteristic is classified separately, since such a process would take a considerable time. Secondly, the parallel strategy may be the more accurate, since in serial classification information about each parameter must be stored until the last characteristic has been classified, before a response to the stimulus as a whole may be executed. The earlier information in the series may be affected by factors which interfere with the stored information and result in an inaccurate perception.

Thanks are due to Dr D. E. Broadbent for his advice, to Mr D. C. V. Simmonds for technical assistance, and to the Royal Navy for supplying subjects and equipment.

References

Neisser, U. (1963*a*). Decision-time without reaction-time: experiments in visual scanning. *Am. J. Psychol.* **76**, 376–85.

Neisser, U. (1963*b*). Multiplicity of thought. *Br. J. Psychol.* **54**, 1–14.

Selfridge, O. G. & Neisser, U. (1960). Pattern recognition by machine. *Scient. Am.* **203** (3), 60–8.

Welford, N. T. (1952). An electronic digital recording machine—the SETAR. *J. sci. Instr.* **29**, 1–4.

(*Manuscript received* 4 *April* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 205–212
Printed in Great Britain

MEANING MEDIATION IN VERBAL TRANSFER

By R. KANUNGO
Dalhousie University, Halifax, Nova Scotia

The role of synonymity and antonymity as mediators in verbal transfer in the absence of any confounding effects of associative connexions was studied. Using an *A–B*, *A–B′*, *A–B* paradigm, subjects learned under three different conditions in each of which *B* and *B′* had least pre-laboratory associative connexions. In a control condition, *B* and *B′* were semantically unrelated and, in the other two experimental conditions, *B* and *B′* had either synonym or antonym relation. Results suggest that implicit meaning reaction alone can mediate verbal learning. Positive and negative transfer effects were obtained in the synonym and antonym condition respectively. Such mediational effects were more pronounced in fast than slow learners.

The importance of similarity in meaning between the elements of two successive paired-associate learning tasks in verbal transfer has been amply demonstrated (Osgood, 1953; Jenkins, 1963). Positive transfer effects have been obtained either using an *A–B*, *A–B′* paradigm in which stimuli are identical and the response terms are different but highly similar in meaning (Bastian, 1961; Haagen, 1943; Runquist & Marshall, 1963; Osgood, 1946; Underwood, 1951) or using an *A–B*, *A′–B* paradigm in which responses are identical and stimuli are different but highly similar (Ryan, 1960). Two different explanations have been offered to account for the underlying processes responsible for these effects. Osgood (1953) suggested that covert meaning reaction can mediate in verbal learning, and transfer effect observed in *A–B*, *A–B′* or *A–B*, *A′–B* paradigms could be explained in terms of mediation of connotative meaning. Words of similar meaning (synonyms) are assumed to be mediated by similar implicit reactions, whereas words of opposite meaning (antonyms) are assumed to be mediated by antagonistic implicit reactions. Thus Osgood (1946, 1948) concluded that, in an *A–B*, *A–B′* paradigm, use of synonyms as responses yields positive transfer through generalization of excitatory tendency, while use of antonyms yields negative transfer through generalization of inhibitory tendency.

This conclusion, however, has recently been questioned by studies where use of antonyms in an *A–B*, *A–B′* paradigm as response items (Bastian, 1961) and in an *A–B*, *A′–B* paradigm as stimulus items (Ryan, 1960) has resulted in positive transfer. These results have suggested the alternative explanation that verbal generalization occurs, not because of mediation of implicit meaning response, but because of pre-laboratory associative connexions between inter-list stimulus or response items. Thus Bastian (1961) suggested that earlier studies dealing with verbal transfer have confounded the effects of two often correlated but separate variables: the degree of semantic similarity and the degree of associative connexions. Separating these two variables, Bastian found that verbal generalization from *A–B* to *A–B′* is fundamentally mediated by pre-laboratory associative connexions between *B–B′* rather than by any semantic relation between them. One may assume quite readily, however, that synonyms are as much a part of the meaning reaction as antonyms (Deese, 1962). Meaning of a word refers to a common state or dimension

with synonyms and antonyms as bipolar parts. This assumption questions the validity of the conclusion that meaning response unsupported by associative connexions does not mediate in verbal transfer. It may be argued that positive transfer effects can result through generalization of excitatory tendency across paired-associate lists containing either antonyms or synonyms in analogous inter-list positions. For both Bastian's (1961) and Ryan's (1960) studies, some positive transfer was obtained when words in analogous inter-list positions were synonyms but maximum positive transfer was obtained when words in analogous inter-list positions were linked by two factors: antonym relation and associative connexion. While these studies suggest that pre-laboratory associative connexions do mediate verbal transfer, the results were not sufficient to preclude the possibility that mediation in terms of meaning does not occur. This calls for further research on the two variables—semantic relations (synonymity and antonymity) and associative connexion—in isolation, in order to determine their respective roles as mediators in verbal transfer.

Recently, Schwenn & Underwood (1965) studied the role of associative connexions as a mediator using a simulated A–B, A–B' transfer paradigm. They tried to eliminate any confounding effects of semantic relations. But as yet no attempt has been made to study the role of semantic relations as a mediator in the absence of any confounding effects of associative connexion. The present study attempted to do this.

Using an A–B, A–B', A–B paradigm, subjects learned under three different conditions in each of which B and B' had least pre-laboratory associative connexion. In a control condition, B and B' had no semantic relation, and in the other two conditions B and B' had either synonym or antonym relations. Differential predictions stemming from different theoretical positions were tested. If transfer effects occur only because of associative connexions and not because of semantic relations between inter-list responses, then no significant difference in the transfer or retroaction effects would be observed. Any difference obtained among the three conditions would indicate that semantic relations do mediate verbal transfer. In this case there are two possibilities. If synonymity and antonymity equally represent meaning response, then both synonym and antonym conditions would yield equivalent amounts of positive transfer and retroactive facilitation when compared with the control. If, on the other hand, synonyms of a word represent similar meaning responses and its antonyms represent different or even opposed meaning responses (Osgood, 1953), then only the synonym condition would yield positive transfer as compared with both the control and the antonym conditions. The antonym condition either may not differ from the control or may yield negative transfer and retroactive interference.

METHOD

Materials and apparatus

The experiment made use of three conditions of transfer: response synonymity, response antonymity and control. In each condition there were three stages: learning an original set of responses, learning another set of responses where transfer can be observed and relearning the original responses to observe retroactive effects. Thus the design required four sets of response words: an original set of responses, a set of synonyms of the original responses, a set of antonyms of the original responses and a set of control words having no semantic relation with the original responses. In addition, the design also required that associative connexions between the original

responses and their synonyms or their antonyms or the control words should be minimal. To achieve this objective, forty frequently used words were chosen (Thorndike & Lorge, 1944). For each of these words a synonym and an antonym were selected. Thus there were three sets of forty words in each: an original, a synonym and an antonym set. The original set of words was then given to a group of eighty subjects to obtain norms for their pre-laboratory associative connexions. For each word, the subjects were instructed to give as many word associations as they could during 60 sec periods. Analysis of these associations provided a normative measure of the frequency with which a word of the original set evoked the corresponding word in the antonym or the synonym sets. Another group of ninety subjects compared each word of the original set with its antonym or its synonym on a nine-point bipolar similar–dissimilar scale. This provided a norm for the degree of antonymity and synonymity between the sets of words.

Table 1. *Stimulus and response items of the four lists, together with frequency counts, and percentage association and degree of semantic relation between inter-list responses*

(The first of the three List 2 responses for each stimulus belongs to List 2S, the second one belongs to List 2A, and the third one belongs to List 2C. The association values of stimuli and the frequency counts of responses are given in parentheses against each item.)

Common stimulus items	List 1 responses	List 2 responses	Degree of synonymity or antonymity ($n = 90$) Mean	Var.	Percentage association ($n = 80$)
ZEG (25)	Win (AA)	Success (AA)	2·50	0·89	2·5
		Failure (48)	2·82	0·92	0·0
		Canal (48)	—	—	0·0
YUV (26)	High (AA)	Top (AA)	2·12	0·74	2·5
		Bottom (A)	2·38	0·78	0·0
		Daily (AA)	—	—	0·0
JEX (23)	Man (AA)	Male (34)	3·37	0·54	8·75
		Female (38)	3·47	0·54	1·25
		Hammer (34)	—	—	0·0
GIH (22)	Correct (A)	True (AA)	2·94	0·77	2·5
		False (A)	3·67	0·24	2·5
		Allow (AA)	—	—	0·0
QAP (24)	Dark (AA)	Gloom (18)	2·36	0·67	0·0
		Glare (26)	2·37	1·26	0·0
		Mold (22)	—	—	0·0
VAF (22)	King (AA)	Master (AA)	2·18	0·89	0·0
		Servant (A)	3·18	0·98	5·0
		Print (A)	—	—	0·0
WUB (21)	Favor (AA)	Support (AA)	2·42	0·87	0·0
		Oppose (A)	3·30	0·43	0·0
		Slow (A)	—	—	0·0
BOF (31)	Gentle (AA)	Tame (25)	2·08	0·79	0·0
		Harsh (17)	3·30	0·41	3·75
		Parcel (23)	—	—	0·0

On the basis of these norms, four lists of paired associates, one original (List 1) and three transfer lists (Lists 2A, 2S, 2C), were prepared. Each list had the same set of eight CVC trigrams with association values ranging from 21 to 31 (Archer, 1960) as stimulus elements. Response items of List 1 were chosen from the original set of forty words. In List 2A, the response items were antonyms of List 1 responses and in List 2S, the response items were synonyms of List 1 responses. Associative connexion between responses of List 1 and List 2A or 2S was minimal, as indicated by the norms obtained. The response items of List 2C were neither associatively nor semantically related to List 1 responses. The degree of synonymity of List 2S responses and the degree of antonymity of List 2A responses to List 1 responses were kept equivalent. In addition, the three sets of responses of the transfer lists were matched as far as possible for

their frequency of occurrence (Thorndike & Lorge, 1944). The stimulus and response items of all the four lists are presented in Table 1.

For each list five different orders of presentation of the pairs were used to minimize the effects of serial learning. The lists were presented by the anticipation method on a Stowe electronic memory drum at a 2:2 rate with 4 sec inter-trial interval.

Subjects

Forty-five undergraduate students participated in this experiment as a part of their course requirements.

Procedure

All forty-five subjects were tested individually in three different stages: original list learning, transfer list learning, and relearning the original list. Approximately 60 sec elapsed between the successive stages of learning. During the original list learning, all subjects were presented with List 1 on a memory drum. They were given standard instructions for the anticipation procedure. All subjects learned the list to a criterion of two consecutive successful anticipations.

During the second stage of transfer the subjects were divided into three groups of fifteen subjects each, matched as far as possible on the basis of the trial and error scores for learning List 1. One group learned List 2A (antonym group), another group learned List 2S (synonym group), and the third group learned List 2C (control group) during this stage. For each group the learning procedure and the learning criterion were the same as those described earlier for the learning of List 1.

During the final relearning stage all forty-five subjects again relearned List 1 to a criterion of two consecutive successful anticipations.

RESULTS

Original list learning

Analysis of variance performed on both trial and error scores of the three groups of subjects during original learning revealed that the attempt to equate all three groups on the basis of List 1 learning had been successful. In the case of both trial and error scores, F ratios were less than 1 (D.F. = 2, 42). The mean learning trials and errors to criterion for each group are presented in Table 2.

Transfer list learning

Analysis of variance involving matched groups was performed on transfer list learning scores. Comparison of the three groups revealed that, with respect to the total number of errors during transfer list learning, there was a significant difference among the three groups ($F = 13{\cdot}686$, D.F. = 2, 28, $P < 0{\cdot}001$). The mean number of errors in each of the three groups is presented in Table 2. The synonym group showed a significantly smaller number of errors when compared with both the antonym ($t = 5{\cdot}10$; D.F. = 28; $P < 0{\cdot}001$) and the control ($t = 3{\cdot}55$; D.F. = 28; $P < 0{\cdot}01$) groups. No significant difference was observed between the antonym and the control groups ($t = 1{\cdot}56$). With respect to the total number of trials to reach the criterion, again a significant difference between the three groups was observed ($F = 3{\cdot}57$; D.F. = 2, 28; $P < 0{\cdot}05$). The mean number of trials in each of the three groups as presented in Table 2 reveals a pattern similar to the mean number of errors. The performance of the synonym group was significantly better than the performance of the antonym group ($t = 2{\cdot}68$; D.F. = 28; $P < 0{\cdot}01$). The performance of the control groups, however, was not significantly different from either the synonym ($t = 1{\cdot}27$) or the antonym ($t = 1{\cdot}40$) group.

Recall of original list responses

The total number of correct anticipations of List 1 responses during the first relearning trial were analysed for the three groups to obtain a measure of retroactive inhibition. The mean number of correct anticipations in each group is presented in Table 2. Significant differences among the groups were indicated by an F ratio of 4·45 (D.F. = 2, 28; $P < 0{\cdot}05$). Further analysis revealed that, while the synonym group recalled better than both the antonym ($t = 2{\cdot}82$; $P < 0{\cdot}01$) and the control ($t = 2{\cdot}28$; $P < 0{\cdot}05$), the antonym and the control groups did not differ significantly from each other ($t = 0{\cdot}54$).

Table 2. *Mean trials and errors to criterion for the three groups during learning successive lists*

	Learning original list		Learning transfer list		Recall of original list responses	Relearning original list	
	Trails	Errors	Trials	Errors		Trials	Errors
Synonym group							
Mean	12·60	47·13	10·73	30·87	5·07	5·60	8·13
S.D.	2·73	18·78	2·72	11·39	2·23	2·33	7·14
Antonym group							
Mean	12·67	47·60	12·40	45·07	3·67	7·13	13·80
S.D.	2·56	14·28	1·36	7·42	0·79	1·26	4·49
Control group							
Mean	12·33	46·13	11·53	40·73	3·93	5·73	10·07
S.D.	2·68	16·36	1·15	4·78	1·01	0·73	3·53

Relearning of the original list

Another measure of retroactive inhibition was obtained by analysing performance during relearning of the original list. The results of the analyses of variance of the total number of errors and of the total number of trials to criterion indicated significant differences among the groups ($F = 4{\cdot}64$; D.F. = 2, 28; $P < 0{\cdot}05$; $F = 6{\cdot}31$; D.F. = 2, 28; $P < 0{\cdot}01$—for error and trial scores respectively). The means of errors and trials for each of the three groups during relearning List 1 are also presented in Table 2. Comparison of these means revealed that the antonym group took significantly more trials to reach the criterion of relearning List 1 as compared with both the synonym ($t = 3{\cdot}21$; D.F. = 28; $P < 0{\cdot}01$) and the control ($t = 2{\cdot}93$; D.F. = 28; $P < 0{\cdot}01$) groups. Likewise, the antonym group committed more errors than the synonym group ($t = 2{\cdot}99$; D.F. = 28; $P < 0{\cdot}01$). The former group also committed more errors than the control group, but this difference only approaches significance ($t = 1{\cdot}97$; D.F. = 28; $P < 0{\cdot}10 > 0{\cdot}05$).

A closer look at the standard deviations presented in Table 2 reveals that, while the variances of the three groups were similar during List 1 learning, they tended to differ during List 2 learning and relearning List 1. Thus there was a need to check if such heterogeneity of variance had affected the significance level of results presented earlier. In order to attain greater homogeneity of error variance, the trial and error scores were transformed by using the formula: Score on List 2 (or relearning List 1) divided by score on List 1 multiplied by 100. For the recall scores on the first

relearning trial, square-root transformation was used. Analyses of variances performed on these transformed scores yielded results essentially similar to those presented above, except for raising the significance level of the results in a few cases.

There is some compelling evidence (see Table 2) to suggest that the introduction of experimental treatments not only affected the mean learning scores of the three groups, but also affected their variances. The control and the antonym groups showed less variability than the synonym group during both List 2 learning and relearning of List 1. Further inspection of the data revealed that the reason for the differences in the variability lies in the differential effects of the experimental treatment on the slow and fast learners within each group. By separating slow and fast learners within each group on the basis of the median List 1 learning scores, it was observed

Table 3. *Mean learning scores of fast and slow learners*

(Eight slow and seven fast learners in the synonym group, and seven slow and eight fast learners were in each of the other two groups.)

	Learning ability	Trials			Errors		
		Synonym	Antonym	Control	Synonym	Antonym	Control
List 1 learning	Slow	14·75	15·00	14·71	61·37	60·0	59·43
	Fast	10·14	10·62	10·25	30·86	36·75	34·5
List 2 learning	Slow	11·87	12·29	11·86	35·62	45·86	41·71
	Fast	9·43	12·5	11·25	25·43	44·37	39·87
Relearning List 1	Slow	5·62	6·86	5·71	8·62	12·28	10·00
	Fast	5·57	7·37	5·75	7·57	15·12	10·12
		Correct anticipations					
Recall on first relearning trial	Slow	4·75	4·00	4·14			
	Fast	5·43	3·37	3·75			

that the slow learners were relatively unaffected by experimental treatments as compared with the fast learners. This will be seen by comparing the mean learning scores of fast and slow learners as presented in Table 3. For instance, List 1 learning scores reveal that the amount of difference between the fast and slow leaners in each of the three groups is somewhat constant. This was borne out by insignificant interaction effects in 2×3 analyses of variance performed on trial as well as error scores. Analyses of List 2 learning scores revealed significant interaction effects ($F = 3{\cdot}83$; D.F. $= 2, 39$; $P < 0{\cdot}05$; $F = 5{\cdot}97$; D.F. $= 2, 39$; $P < 0{\cdot}01$—for trials and errors respectively) suggesting that the difference between scores of slow and fast learners in the synonym group is greater than those in the antonym or the control group. This is why greater variances were observed in the synonym group when compared with either the antonym or the control group. Furthermore, the mean scores of the fast and slow learners in Table 3 clearly suggest greater variability in fast learners than slow learners resulting from experimental treatments. In other words, fast learners are more affected by meaning mediation than slow learners. Similar trends are also observed in the case of recalling and relearning List 1 (see Table 3), although analyses of variance performed on these data resulted in non-significant interaction effects.

Inter-list intrusions

An inspection of inter-list intrusions reveals that, during learning of the transfer lists, the total number of intrusions of List 1 responses were 26, 18 and 11 respectively for the antonym, the synonym and the control groups. The total number of subjects showing such intrusions were thirteen, nine and seven respectively. A similar trend was noticed during relearning List 1. The total numbers of intrusions of transfer list responses were twenty-three, twenty and fourteen, and the total numbers of subjects showing these intrusions were thirteen, seven and nine respectively for the antonym, the synonym and the control groups. It will be noticed that compared with the synonym and the control groups, a larger number of subjects in the antonym group showed overt inter-list intrusions during transfer list learning as well as during relearning List 1 (in both the cases $\chi^2 = 5{\cdot}43$; D.F. $= 2$; $P < 0{\cdot}10 > 0{\cdot}05$). This suggests a tendency for greater response competition when the response members of the successive lists have opposed meaning relations than when they have either same or different meaning. Further, a closer examination of intrusions in the synonym and the control groups revealed that, for those subjects who showed overt intrusions, similarity of response meaning in the synonym group caused greater interference (nine and seven subjects had a total of eighteen and twenty intrusions respectively during transfer list learning and List 1 relearning stages), than the control (seven and nine subjects had a total of eleven and fourteen intrusions respectively during transfer list learning and List 1 relearning stages).

Discussion

The results clearly indicate that, eliminating any confounding effects of associative inter-list relation, semantic relations alone can mediate verbal transfer. Therefore the suggestion that 'generalization between words judged to be similar in meaning occurs *only because* associative tendencies connecting them are usually present in subjects' verbal behaviour' (Bastian, 1961, p. 78) has to be rejected. It is, however, recognized that pre-laboratory associative relations between words exert a potent influence on verbal transfer and that this overshadows the effects of semantic relations between words. This would appear to be the case in Bastian's (1961) and Ryan's (1960) studies.

The pattern of trial and error scores during learning of the transfer lists and the recall of List 1 responses during the first relearning trial seem to suggest that, as compared with the control group, positive transfer effect and retroactive facilitation were obtained in the synonym group. On the other hand, the pattern of intrusion data and the trial and error scores during relearning List 1 seem to suggest negative transfer effect and retroactive interference in the antonym group. At this point, two questions may be raised. First, why the synonym group when compared with the control did not show significant positive transfer during relearning List 1 while it did show such transfer during learning the transfer list. Secondly, why the antonym group when compared with the control did not show significant negative transfer during the transfer list learning while it showed such transfer during relearning List 1. Two factors might account for these results. First, the relearning of List 1

was relatively much easier, so that differences between the groups, particularly the synonym and control groups who did not have to face response competition with opposed meaning responses, were minimized. This probably accounts for the lack of significant difference between the synonym and the control groups during the relearning stage. Secondly, the lack of significant difference between the antonym and the control groups during the transfer list learning stage was presumably due to the relative ease of recognizing a synonym relation rather than an antonym relation among inter-list response words. Answers to post-experimental queries revealed that during transfer list learning subjects in the antonym group took a longer time to recognize the responses in their transfer list as antonyms of and not just different from the corresponding responses in the original list. Subjects in the synonym group, however, recognized the responses of their transfer list as synonyms much earlier during learning.

These results are in line with the predictions derived from Osgood's (1953) suggestion that implicit connotative meaning reactions can mediate verbal learning. The study gives considerable support to Osgood's theory of meaning in demonstrating opposite effects of synonym and antonym relations in verbal transfer. Finally, these findings suggest that meaning mediation in verbal transfer interacts with initial learning ability. Mediational effects seem to be more pronounced in fast than in slow learners.

This work was supported in part by the Canadian Defense Research Board, Grant 9401–22, and in part by the University Research Development Fund, Grant X-84-80. The author is indebted to Bruce Earhard for his many helpful comments, and to Don Moors and Lynn Ross for their assistance in collecting the data.

REFERENCES

ARCHER, E. J. (1960). Re-evaluation of meaningfulness of all possible CVC trigrams. *Psychol. Monogr.* **74**, No. 10 (whole no. 497).

BASTIAN, J. (1961). Associative factors in verbal transfer. *J. exp. Psychol.* **62**, 70–9.

DEESE, J. E. (1962). On the structure of associative meaning. *Psychol. Rev.* **69**, 161–75.

HAAGEN, C. H. (1943). Learning and retention as a function of the synonymity of original and interpolated tasks. (Unpublished doctoral dissertation, State University of Iowa.)

JENKINS, J. J. (1963). Mediated associations: Paradigms and situations. In *Verbal Behaviour and Learning*, ed. C. N. Cofer and B. Musgrave. New York: McGraw-Hill.

OSGOOD, C. E. (1946). Meaningful similarity and interference in learning. *J. exp. Psychol.* **36**, 299–301.

OSGOOD, C. E. (1948). An investigation into the causes of retroactive interference. *J. exp. Psychol.* **38**, 132–54.

OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology*. New York: Oxford University Press.

RUNQUIST, W. N. & MARSHALL, M. A. (1963). Transfer, synonymity, and anticipatory interval in paired-associate verbal learning. *Am. J. Psychol.* **76**, 281–6.

RYAN, J. J. (1960). Comparison of verbal response transfer mediated by meaningfully similar and associated stimuli. *J. exp. Psychol.* **60**, 408–15.

SCHWENN, E. & UNDERWOOD, B. J. (1965). Simulated similarity and mediation time in transfer. *J. verb. Learn. verb. Behav.* **4**, 476–83.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Columbia University Press.

UNDERWOOD, B. J. (1951). Associative transfer in verbal learning as a function of response similarity and degree of first list learning. *J. exp. Psychol.* **42**, 44–53.

(*Manuscript received* 17 *August* 1966)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 213–218
Printed in Great Britain

REMEMBERING OF WORDS BY PSYCHOTIC AND SUBNORMAL CHILDREN

By BEATE HERMELIN and N. O'CONNOR
Institute of Psychiatry, Medical Research Council, Maudsley Hospital, London

It was predicted that recall of structured and unstructured verbal material would differ less in a group of psychotic children than in subnormal controls. This hypothesis was confirmed. In addition, psychotic children had better recall scores than the controls.

Studies on psychological functions of psychotic children show agreement in regard to their special disabilities on verbal tests (Rutter, 1966; Gillies, 1965; Goldfarb, 1961). In most of the samples of psychotic children which have been investigated, about half have been mute (Kanner & Eisenberg, 1955; Eisenberg, 1956; Rutter, 1966; Mittler, Gillies & Jukes, 1966). In Rutter's series only ten out of sixty-three cases obtained a normal level of speech development.

If speech does develop it usually has some peculiar characteristics such as extreme concreteness, specificity, pronominal reversal and echolalia. Goldstein (1959) noted that all that the children seem to be able to do with language is remember and repeat rather than understand it. A question to the child such as 'Do you want a sweet?' tends to be stored by him without being analysed and later emerges unchanged in the analogous stimulus situation. Repetition of a verbal sequence implies the ability to remember it. Confirming this, Rutter (1966) reported that immediate tests were relatively well performed by his sample of autistic children. Tubbs (1966) found verbal rote memory the only item from a psycholinguistic test on which psychotic children approached a near normal level.

Miller & Selfridge (1950) state that 'communicative behaviour, perhaps more than any of man's other activities, depends on patterning for its significance and usefulness'. If analysis and patterning were absent in speech as well as in language comprehension of psychotic children, this might explain their inability to communicate, without the necessity of evoking concepts such as 'social and emotional withdrawal'. Goldfarb (1961) reported that normal children under conditions of delayed auditory feedback of their own speech showed more disturbance in their speech patterns than did psychotic children. This seems to indicate that speech in psychotic children is less dependent on monitoring. Such an interpretation, rather than one of lack of hearing of the delayed feedback, was suggested by the greater non-verbal confusion, restlessness and distress shown by the psychotics than by the normals under those conditions.

Analysis and patterning of language most usefully conforms to verbal meaning and contextural structure. This can be defined by the statistical properties of speech sequences, i.e. the degree to which the choice of a particular word depends upon the words that preceded it. Word sequences which are completely random are termed zero approximations to language, and the order of approximation increases with the amount of contextual restraint. Verbal recall in normals is better for material with a higher than with a lower order of approximation to the common language.

Frequency of word usage also affects recall, which increases with word frequency.

However, Miller (1956) has clearly shown that the length of the memory span is limited by the number of items rather than by the amount of information per item. Deese (1961) has suggested that the improved recall for structured or familiar sequences of words is due not so much to the greater amount retained as to the subject's ability to reconstruct the material by associating or guessing on the basis of his knowledge of the characteristics of the language. The following experiments compare the recall ability of subnormal and psychotic children. They also investigate the relationship of immediate verbal recall with word frequency, sequential probability and conceptual patterning. It was predicted that recall of structured and unstructured material would differ less in the psychotic group than in the controls.

EXPERIMENT I

Method

Subjects

The subjects were twelve psychotic and twelve severely subnormal children aged between eight years and fourteen years five months with a mean age of ten years eight months. They were matched on the Peabody Picture Vocabulary Test for the visual recognition of object names. Mental ages on this test ranged from MA 2–6 to MA 10–8; mean MA 4–3. The subjects from the two groups were also individually matched on their immediate memory span for digits, which ranged from 2 to 7, with a mean span of 4·5 digits.

The diagnosis of infantile autism or childhood psychosis was in all cases made by two psychiatrists; in addition, all psychotic children showed a minimum of four of Creak's (1961) nine diagnostic features. In all cases this included retardation or abnormalities of speech. Children who were completely speechless were excluded from the sample, as were those with any obvious physiological abnormalities. However, no results of detailed examinations of EEG, or neurological and biochemical dysfunctions were available, and the selection was made on a behavioural basis. Half of the children lived at home and half in residential institutions. The controls were a heterogeneous group of severely subnormal institutional children, from which those with mongolism, epilepsy or any other obvious sensorimotor dysfunction were excluded.

Material

This consisted of sequences of one-syllable words, selected from the vocabulary list of words used by severely subnormal subjects (Mein, 1961). This list contains 1720 words and their frequency of occurrence in the speech of a sample of eighty cases.

The selected sequences ranged from three to eight words per message: six sequences of three, six of four, six of five, six of six, six of seven and six of eight words. These words were arranged in different ways for different sets of material. Set A contained simple English sentences. Set A^1 contained the same words as set A but arranged in random order, so that the messages resembled zero-order approximations to English. These arrangements were duplicated with different words of the same frequency for sets B and B^1. There were thus four sets of words, each containing thirty-six messages, all matched for average word frequency. Each single message, whether of three, four, five, six, seven or eight words and whether randomly arranged or as a sentence, had an average word frequency between 63·4 and 65·7 per cent.

In addition, there was a set of 36 one-syllable word sentences, which varied in length from three to eight words, six for each block, and which had an average word frequency of 34 per cent (C). No randomization of these words was prepared. Examples of the words are given in Table 1.

Procedure

Each subject was presented with three types of word sequences for recall. These were sentences of frequent words (A or B), the words from one of these two sets arranged in randomized order (A^1 or B^1) and sentences of less frequent words (C). Those who were presented with the sentences in set A were given the randomized series B^1 and those who had to recall set B sentences were presented with the randomized sequences (A^1). All subjects were given set C. Thus each subject

was required to recall words matched for frequency in contextual and random order, as well as less frequent words in contextual order. The messages were presented vocally at the rate of two words per second, and the order in which different types of sequences were presented was varied randomly between subjects, according to a balanced design. Within each set, six three-word messages had to be recalled first, followed by six four-, five-, six-, seven- and eight-word messages. Testing was continued until a complete block of six messages of a given length failed to be recalled correctly. Recall was immediate and unpaced, and a correct response was one in which all the items from a message were recalled in the correct order.

Table 1. *Examples of material*

Frequent word sentences (A)

Do not run
We went to town
The bird builds its nest
When do you go to school
I saw a car by the house
The big ship sails on the blue sea

Frequent word sentences (B)

Play with me
What is the time
We all say good night
The fish swims in the pond
Two men rode past the old church
The nurse has come to fetch the boy

Frequent random words (A^1)

Now [illegible] here
Some [illegible] a went
Home give down red ball
By is go tree stroke lets
Do lamps you the up girl one
Him blue the cat tea give a go

Frequent random words (B^1)

Comb take play
Light what leaf the
The at down a big
Cold to next is have time
Milk score the is a on church
Want fetch come go film has do say

Infrequent word sentences (C)

Let us swap
Shade this young plant
Who has my short thread
Does he want to sleep late
You may count nine more square beads
This black sack was found near the shore

Results

Results were scored in terms of the number of messages correctly recalled for any given length of message and type of material. The total scores for groups and conditions are given in Table 2.

Three analyses of variance were carried out. The difference scores between types of material were compared for the groups. For these analyses scores of three- and four-word messages only were used, as too many scores of 0 were obtained for longer messages. No difference in recall scores between sentences of more frequent and sentences of less frequent words was found, but psychotics had higher recall scores than subnormals under both conditions ($P = 0{\cdot}025$). Likewise, comparing recall for sentences of infrequent words with random arrangements of frequent words did not result in a significant conditions term. Psychotic children did better than the controls in both instances ($P = 0{\cdot}001$). An analysis of the difference between scores from sentences

and randomly arranged words matched for frequency resulted in a groups by conditions interaction ($P = 0{\cdot}025$). While the subnormal controls did significantly better with sentences than with random sequences, no significant difference was found with the psychotics.

From these results it can be concluded that psychotic children had better immediate recall scores than subnormals. This was particularly evident when the material consisted of sentences containing infrequent words or random word sequences. It is also clear that the effect of structure and constraint on recall was less marked for psychotic children than for subnormal controls.

Table 2. *Number of messages correct* (*optimum* = 1252)

	Psychotics	Subnormals
Frequent word sentence	989	928
Rare word sentence	939	853
Frequent random words	896	788

EXPERIMENT II

As contextual restraint affects recall in psychotic less than in subnormal children, it might be inferred that the former remain relatively unaware of the different degree of structure in the two types of verbal material. Deese (1961) has pointed out that the advantage in recalling structural material is accounted for by the active process of recoding, and by the prediction of sequential probabilities. A further experiment investigated the degree of active coding and categorizing of words by psychotic children. In this study word sequences in excess of the memory span were presented and free recall was required. Which of the items would be recalled would depend partly on the degree of association between them, as would the rearrangement of serial order in recall. Bousfield (1953) carried out an investigation on normals, presenting in random order a series of words falling into a limited number of categories. It was found that related items tended to be recalled together in clusters, and the degree of such clustering and amount of recall were correlated. This suggests that the presentation of individual words leads to active coding and categorizing on the part of the subject. As the interpretation of the results of the first experiment has been in terms of a lack of such coding processes in psychotic children, the following experiment tested this hypothesis more directly.

Method

Four word sequences, each eight words long, were presented to the same subjects who had taken part in the first experiment. As the immediate memory span for digits ranged from 2 to 7 (mean 4·5), this was beyond the memory span of all subjects. The four sequences were given to each subject, and were presented in a random order between subjects according to a balanced design. Presentation rate was at two words per second and instructions were to reproduce any remembered item regardless of order.

Material and procedure

The four word sequences are given in Table 3. They consisted of one-syllable words and were matched for mean word frequency, which was 57 per cent for each message. It will be seen that sequence 1 consisted of two categories of words, colours and numbers; sequence 2 was a short sentence with digits interpolated between the words; in sequence 3 names of table utensils were

presented together with interspersed random words, and sequence 4 contained eight random words as a control condition. Thus sequences 1 and 2 lent themselves to clustering into two categories of four items each, and sequence 2 contained a sentence which might be regarded as a single chunk of information. Sequence 3 contained words from one category and sequence 4 contained none.

Table 3. *Material for Expt II*

(1)	Blue	three	red	five	six	white	green	eight
(2)	Nine	this	one	tea	four	is	ten	cold
(3)	Glass	hand	cow	pot	cup	meat	spoon	place
(4)	Ship	gun	night	home	jug	time	lamp	farm

Results

A first analysis of variance, using the number of items correctly recalled, showed that neither the group term nor the conditions term nor interaction approached significance. Thus, in terms of items recalled, the groups were similar and comparable. A further analysis compared the amount of clustering in items recalled. Clustering was defined as any reorganization of serial order by which related items which had been presented interspersed with other items were presented together. A sequence of two such items was scored 2, a sequence of three, 3, and a sequence of four, 4. This analysis, using sequences 1, 2 and 3, did not demonstrate a different degree of clustering for the different sequences, but did show that significantly more clustering of related items occurred in the subnormal than in the psychotic group ($P = 0{\cdot}05$). A χ^2 test comparing the number of instances in which the material was rearranged into categories by the two groups confirmed this result. In twenty-one out of thirty-six possible instances the severely subnormal children recoded the words in this way, while this occurred only nine out of thirty-six times in the psychotic group ($P = 0{\cdot}01$).

Discussion

The results from these two experiments confirm the hypothesis that psychotic children at least equal severely subnormal ones in their immediate recall ability of verbal material. However, while sequential probabilities and conceptual categories affect the recall of the subnormal groups, this occurs to a significantly lesser extent with the psychotics. If one accepts that patterning and coding of words occurs in accordance with previous learning and experience with the language and the resulting expectations, one might infer that the psychotic children show little evidence of such previous learning. This conclusion applies to statistical as well as to conceptual associations. While the first experiment used word sequences which were ordered in terms of their sequential probabilities, the second experiment used verbal material which could be ordered in terms of conceptual categories. The psychotic children made less use of either of these two modes of patterning than did the controls. This illustrates the difference between their linguistic skill and their knowledge of and memory for words. The groups were matched for verbal mental age on a test concerned with the naming of objects, as well as on immediate memory span for digits. If anything, the psychotics have better recall scores for words than the controls. It thus seems that the verbal deficits of psychotic children, their echolalia and their

inability to communicate verbally are less concerned with the knowledge of words or with emotional withdrawal than with an inability to pattern these words according to their statistical and conceptual properties.

We would like to thank Mrs Sybil Elgar of The Society's School, Ealing; Dr G. O'Gorman of Smith Hospital, Henley, and Dr J. Gibson of St Lawrence's Hospital, Caterham, for the facilities which they made available.

REFERENCES

BOUSFIELD, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.* **49**, 229–40.

CREAK, M. *et al.* (1961). Schizophrenic syndrome in children. *Lancet* **2**, 818.

DEESE, J. (1961). From the isolated verbal unit to connected discourse. In *Verbal Learning and Verbal Behavior*, ed. C. N. Cofer, pp. 11–41. New York: McGraw-Hill.

EISENBERG, L. (1956). The autistic child in adolescence. *Am. J. Psychiat.* **112**, 607–12.

GILLIES, S. (1965). Some abilities of psychotic children and subnormal controls. *J. ment. Defic. Res.* **9**, 89–101.

GOLDFARB, W. G. (1961). *Childhood Schizophrenia.* Cambridge, Mass.: Harvard University Press.

GOLDSTEIN, K. (1959). Abnormal mental conditions in infancy. *J. nerv. ment. Dis.* **128**, 538–57.

KANNER, L. & EISENBERG, L. (1955). Notes on the follow-up studies of autistic children. In *Psychopathology in Childhood*, ed. P. H. Hoch and J. Zubin, pp. 227–39. New York: Grune & Stratton.

MEIN, R. (1961). *A List of Words used in Conversation of Severely Subnormal Patients.* Cell Barnes and Harperbury Group Hospital Management Committee.

MILLER, G. A. (1956). The magical number seven, plus or minus two. *Psychol. Rev.* **63**, 81–97.

MILLER, G. A. & SELFRIDGE, J. A. (1950). Verbal context and the recall of meaningful material. *Am. J. Psychol.* **63**, 176–85.

MITTLER, P., GILLIES, S. & JUKES, E. (1966). Prognosis in psychotic children: Report of a follow-up study. *J. ment. Defic. Res.* **10**, 73–83.

RUTTER, M. (1966). Psychotic children in adolescence and early adult life. In *Early Childhood Autism*, ed. J. Wing. London: Pergamon.

TUBBS, V. K. (1966). Types of linguistic disability in psychotic children. *J. ment. Defic. Res.* **10**, 230–40.

(*Manuscript received* 5 *October* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 219–222
Printed in Great Britain

MEMORY AFTER READING ALOUD AND READING SILENTLY

By E. C. POULTON and C. H. BROWN
Medical Research Council, Applied Psychology Research Unit, Cambridge

Twenty-four housewives read passages aloud and silently in balanced order, and then answered questions on them. Time allowed for silent reading was matched to time taken to read aloud. Memory for the first 30 per cent of the passage was reliably ($P < 0{\cdot}02$) less after reading aloud, as if the monitoring of speech put an extra load upon the central mechanisms, thus interfering with memorization. For most of the remainder of the passage memory was equally good after reading aloud and silently. Presumably the vocal output had been programmed to run with the minimum of attention. The last 10 per cent of the passage was remembered reliably better after reading aloud ($P < 0{\cdot}01$), as if vocalization had given it extra emphasis.

When a housewife reads silently to herself and knows that she will have to answer questions on what she is reading, she has the task of selecting and storing information so that she can recover it to answer the questions that she will be given. But if she has also to read the material aloud to an audience, she has a dual task: she has to communicate with her audience as well as select and store information. Unlike most dual tasks, in which each task acts as a distractor for the other (see Poulton, 1966), the two tasks in reading aloud are related, in that understanding the passage is important both in communicating and in selecting and storing information. Thus the effect of distraction, if present, should be less marked.

In the present experiment a measure of 'comprehension' (more correctly described as a measure of memory) was used to assess the degree of distraction. The measure may not distinguish between the dual and single tasks if in performing the dual task the housewife concentrates upon storing information and employs only her spare mental capacity in communicating. However, this is not likely to happen when, as here, she has to read aloud into a microphone with the experimenter listening, and knows that the accuracy of her tape-recording will be checked later.

Recently Murray (1965) and Woodhead (1966) have shown that, when lists of consonants or short words are presented at fast rates, vocalization during presentation assists immediate recall. If a prediction could be made from this rather different experimental task, it would favour reading aloud, despite its dual-task nature.

Method

Material

The four simple passages, each of about 450 words on widespread infectious diseases, were similar to those used previously in experiments on letter design (Poulton, 1965; Poulton & Brown, 1968) and on reading in flickering light (Poulton, Kendall & Thomas, 1966). There were ten open-ended questions on each passage to test for memory. The questions were spread evenly over the text. Both passages and questions were duplicated versions prepared with an Olivetti Lexikon 80 Elite typeface.

Design and subjects

Twenty-four housewives, aged between 21 and 65 years, were allocated in order of arrival to one of three equal groups. One group was instructed to read aloud with expression, the second

to read aloud as quickly as possible, and the third to read aloud so quickly that punctuation, expression and meaning were neglected.

The same two passages were always used in the practice. The housewife had first to study a passage silently for 90 sec if she was in either of the two slower groups, but for only 75 sec if she was in the fastest group. After study she had to answer ten questions on the passage. This was followed by reading the second passage aloud and then answering the corresponding ten questions. Half of the housewives in each group started with one passage, the other half started with the other passage.

The experiment proper was a repeat of the practice, using the other two passages, except that half of the housewives in each group read aloud first, and half read silently first. The time allowed for silent reading was matched to the time taken to read aloud. For the housewives who read silently first, the time allowed was estimated from the time they had taken to read aloud in the practice.

Procedure

The housewife read aloud into a microphone feeding a Ferrograph magnetic tape-recorder. Before reading the practice passage aloud, she was given a comparable prepractice passage to read aloud. After reading it, the tape-recording was played back to her. If she was supposed to be reading quickly but took more than 2 min to read the 450 words, she was asked to read the subsequent practice passage more quickly. Two of the housewives in the fastest group who took longer than 2 min were asked to read the prepractice passage again, this time more quickly. With the third housewife in this group who took longer than 2 min, it was apparent that she was in fact reading as quickly as she could, so she was not asked to re-read the prepractice passage. The recording of the practice passage itself was not played back. Instead the housewife was given the ten questions to answer on it.

Scoring

After the experiment the tape-recordings of the passages read aloud were played back. Reading errors were noted for each tenth of the passage related to a particular test question to determine whether errors on a test question could have been caused by errors in reading aloud. A standardized scoring procedure was used for the answers to the questions.

RESULTS

The mean times taken to read aloud by the three groups were 163, 119 and 110 sec. The corresponding mean times allowed for reading silently were 166, 118 and 111 sec, which were reasonable matches. The mean numbers of misreadings on reading aloud were 1·7, 3·0 and 2·0 respectively. The samples were not reliably different on one-tailed Mann–Whitney U-tests (Siegel, 1956). The misreadings were distributed apparently at random throughout the passages. None of the misreadings should have affected performance on the test questions if the housewife remembered what she had said rather than what was written. One housewife reported that she did not quite finish reading the passage silently in the time allowed. Eight said that they finished reading the passage only once, while the remaining fifteen read part or all of the passage a second time.

The number of questions answered correctly was 66 per cent after reading silently and 64 per cent after reading aloud. The difference was not reliable statistically ($P > 0{\cdot}05$). Fig. 1 shows the distribution of correct answers over the ten questions on each of the two passages, one from each tenth of a passage, for all twenty-four housewives reading both silently and aloud. Since the tenths of the passages and the questions on them were not equated for difficulty, only the relative heights of the two functions in the figure necessarily mean very much. In reading aloud, the

storage of information was relatively ineffective over the first three-tenths of the passages, the number of questions answered correctly being 51 per cent, compared with 71 per cent for reading silently. For the first three questions taken together, analysis of variance indicated that the difference in favour of reading silently was reliable at the 0·02 level.

Reading aloud showed a compensatory advantage on the last question; 81 per cent correct answers, compared with 56 per cent correct for reading silently. The difference was reliable at the 0·01 level on a two-tailed Wilcoxon test (Siegel, 1956). Excluding the housewife who reported that she did not reach the end of the passage when reading silently, the difference was still reliable at the 0·02 level. It is not possible to tell whether performance on the last question of each passage represents some form of end-effect, or whether the penultimate question was unusually difficult.

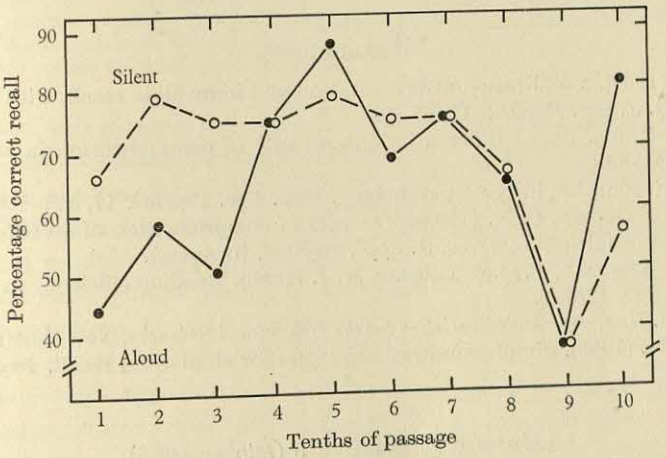


Fig. 1. Mean percentage of correct answers to the ten questions, one on each tenth of the passage. Each point represents data from the same twenty-four housewives.

Discussion

The detrimental effect of reading aloud upon recall of the first part of the passage occurred despite recent practice at reading aloud on two passages, a prepractice and a practice passage. Presumably, when starting to read the test passage aloud, the housewife had to attend to her vocal output, its loudness, pronounciation, etc. She was thus unable to store the information from the first part of the passage as adequately as she could when reading silently. But since reading aloud is a highly over-learnt skill, after the first approximately 150 words she was able to programme her vocal output satisfactorily. It could then continue to run with the minimum of attention, leaving her more free to concentrate upon storing the information in the passage.

Fig. 1 shows that reading aloud was actually advantageous in answering the last question. If the housewife who failed to finish the silent passage in the time allowed is excluded, the result does not appear to be related to the rate of silent reading. Six out of the fifteen housewives who reported reading part or all of the silent passage a second time showed the effect; one showed the reverse effect. Whereas

three out of the eight housewives who reported reading the silent passage just once through showed the effect, none showed the reverse. These proportions are about equal. They do not suggest that success in answering the last question was due to just having finished reading the passage in the time allowed. If it were, reading aloud would clearly be at an advantage, since here the passage was always read once and once only.

Improved memory for the end of the passage after reading aloud may be related to the results of Murray (1965) and Woodhead (1967), which demonstrated the advantage of vocalization upon the immediate recall of lists of consonants and words. Reading the passage aloud may have helped to leave a more durable memory trace of the last part. This would not have been obliterated by subsequent reading aloud, as the memory traces of the earlier parts of the passage may have been.

REFERENCES

MURRAY, D. J. (1965). Vocalization-at-presentation and immediate recall, with varying presentation-rates. *Q. J. exp. Psychol.* **17**, 47–56.

POULTON, E. C. (1965). Letter differentiation and rate of comprehension in reading. *J. appl. Psychol.* **49**, 358–62.

POULTON, E. C. (1966). Engineering psychology. *Ann. Rev. Psychol.* **17**, 177–200.

POULTON, E. C. & BROWN, C. H. (1968). The rate of comprehension of an existing teleprinter output and of possible alternatives. *J. appl. Psychol.* (in press).

POULTON, E. C., KENDALL, P. G. & THOMAS, R. J. (1966). Reading efficiency in flickering light. *Nature, Lond.* **209**, 1267–8.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

WOODHEAD, M. M. (1966). Simple rehearsal strategies for short-term recall. *Psychonom. Sci.* **6**, 385–6.

(*Manuscript received* 5 *October* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 223–226
Printed in Great Britain

STRUCTURE AND CONTENT IN CONCEPT ATTAINMENT: EFFECTS OF VARYING PROBLEM CONTENT

By N. E. WETHERICK

Department of Social Sciences, University of Bradford

Three equivalent forms of a concept-attainment task were employed on the hypothesis that both unfamiliar problem content (Hebrew letters) and relatively familiar problem content (simple verbal descriptions), which however exposed the problems to proactive interference from individual preconceptions, prejudices, etc., would have the effect of reducing the efficiency with which problems were solved by comparison with neutral problem content (English letters). The second part of the hypothesis was supported but not the first. Young, intelligent subjects may operate as easily with unfamiliar as with familiar symbols, though real verbal content significantly increases the difficulty of the task for them; but this possibility requires confirmation by a more fully balanced experiment.

In an article in this *Journal* (Wetherick, 1966) the writer described a paper-and-pencil test of ability to make the kinds of inference required in concept attainment. One of the principal advantages claimed for the type of problem employed was that any number of structurally equivalent problem series may be generated at will. The problems present groups of symbols in which each position in the group represents an attribute and each symbol an attribute value. In the experiments reported the symbols employed were English letters which are highly overlearned, and consequently easy to perceive and hold in short-term memory. English letters have the further advantage that, when used in groups that do not resemble words, they have relatively few associations; problems employing them are consequently unlikely to be subject to serious proactive interference. Such problems are perhaps the nearest it is possible to approach to the ideal of structure without content. They are well adapted for the study of logical operations in their pure form. In real life, however, concept-attainment problems always have significant content, and this content may be unfamiliar, in which case the material may be difficult to perceive accurately and hold in short-term memory; or it may be familiar and the problems may be subject to proactive interference from individual preconceptions, prejudices, and so on as in the thematic material employed by Bruner *et al.* (1956). The experiment tried to evaluate the effect of varying relative familiarity/unfamiliarity of problem content on the accuracy and speed with which the problems were solved.

Method

The tasks. A series of sixteen problems was constructed of which four presented two positive instances, four three positive instances, four one positive instance and two negative and four one positive instance and three negative. All the problems presented instances drawn from a universe of sixteen possible instances exemplifying four attributes, each of which could take one of two values. The solution was in each case a pair of attribute values. Three different forms of the task were prepared; the problems in each form were structurally equivalent and differed only in content. The first form employed English letters as before; B or C occupied the first letter position, D or F the second, G or H the third and J or K the fourth. The second form employed Hebrew letters which were regularly substituted for English. Hebrew letters are not much like English letters nor are they simple shapes. It was hypothesized that by reason of

their unfamiliarity they would be more difficult to perceive accurately and hold in the short-term store, with the result that the problems would be solved less quickly and/or accurately. The third form of the task employed simple verbal material. Sixteen men were each described as either tall or short with blue or grey eyes, black or brown hair and a dark or pale complexion. It was hypothesized that by reason of its familiarity problems employing this kind of material might be subject to proactive interference from individual preconceptions, etc., with the result that these problems too might be solved less quickly and/or accurately.

Procedure. In each form of the task the subject was presented with two sheets. One sheet bore the sixteen instances in vertical column (in the English and Hebrew letter forms the instances were numbered 1–16; in the verbal form they were designated Mr A., Mr B.,. . .Mr Q., in alphabetical order), and the instructions for the task, which included two fully worked examples, were also on this sheet. The other sheet bore the sixteen problems. In the problems the instances were designated by number or name as appropriate; '1. and 2.', 'Mr A. but not Mr C. or Mr D.'. The subject responded by underlining one of eight pairs of attribute values given to him to choose from or by writing in what he thought was the correct pair if none of the eight seemed correct to him. (In fact the eight pairs of values always included the correct pair. The subject was not told this in order to discourage him from trying to work backwards, testing the given pairs one at a time.) The time allowed for each form of the task was 6 min. It is to be emphasized that the three forms were identical except as regards problem content.

Subjects were 48 first-year university students (22 male, 26 female), members of an introductory course in psychology. The subjects were tested in two groups of 24 and were available for only 1 hr. No relevant background data could be obtained.

Experimental plan. In the experiments previously reported, scores on Raven's Progressive Matrices were available for comparison. It was obviously impossible to obtain full scores on this test in the time available and it would in any case have been necessary to use the 'Advanced Matrices' to obtain satisfactory discrimination among subjects at this level. In the circumstances the best that could be done was to set Raven's Progressive Matrices (1956) with a time limit of 15 min. The scores obtained must obviously not be regarded as adequate measures of non-verbal intelligence but they do permit the subjects to be ranked. The group achieved a mean score of 53·25 (S.D. 4·35); three subjects scored full marks despite the time limit.

Using only two groups of subjects it was not possible to balance out order effects between the three forms of the task. It was therefore decided to balance out effects between the English letter and the verbal form, and to isolate the Hebrew letter form by placing it last in order after the Progressive Matrices. Group I did the English letter form first (6 min.), followed by the verbal form (6 min.), the Progressive Matrices (15 min.) and the Hebrew letter form (6 min.). Group II did the first two items in reverse order.

RESULTS

Results are shown in Table 1. Scores in the two groups do not differ significantly except in the verbal form of the task. The problems had not previously been used with female subjects, but no significant sex differences were found in either group. Scores on the English letter and verbal forms of the task were subjected to analysis of variance. The verbal form was found to be significantly more difficult than the English letter form ($P < 0{\cdot}001$) but the form done first was also significantly more difficult than the form done second ($P < 0{\cdot}01$). There was again no significant difference between the groups. The difference between scores on the Hebrew letter form of the task and the English letter form was marginally significant in group I only; but the Hebrew letter form had the advantage of being done last by both groups. Group I scored higher than group II on both the Progressive Matrices and the Hebrew letter form and it seems likely that this group's lower scores on the English letter form may be attributed to the order in which they were required to attempt the tasks.

Rank order correlations were calculated between scores on the Progressive

Matrices and on the various forms of the task. Since there were no significant differences between groups on the matrices, the English letter form and the Hebrew letter form these three correlations were calculated for the combined groups. All three were significant.

Matrices/English letter form $\rho = 0{\cdot}33$ ($P < 0{\cdot}05$)
Matrices/Hebrew letter form $\rho = 0{\cdot}44$ ($P < 0{\cdot}01$)
English letter form/Hebrew letter form $\rho = 0{\cdot}51$ ($P < 0{\cdot}001$)

Scores on the English letter form and the verbal form of the task correlated significantly in both groups (group I: $\rho = 0{\cdot}45$, $P < 0{\cdot}05$; group II: $\rho = 0{\cdot}53$, $P < 0{\cdot}01$) but there were no significant correlations between scores on the matrices and on the verbal form of the task, or between the verbal form and the Hebrew letter form.

Table 1. *Scores of two groups of subjects on Raven's Progressive Matrices* (1956) *and on three forms of the concept-attainment task*

	Raven's matrices (15 min, max. 60)	English letter form (max. 16)	Verbal form (max. 16)	Hebrew letter form (max. 16)
Group I				
Mean	54·08	9·75	8·04	12·08
S.D.	3·55	3·18	3·52	3·76
Group II				
Mean	52·42	10·50	4·25	11·00
S.D.	4·86	3·37	1·64	3·53
Both groups				
Mean	53·25	10·13	—	11·54
S.D.	4·35	3·29	—	3·76

Discussion

It is an open question how far the matrices scores obtained here ranked the subjects in the same order as a fully adequate measure of non-verbal intelligence would have done. If, however, the ranking is accepted, the correlations with the three forms of the task are consistent with common sense. The correlation with the verbal form is not significant (group I: $\rho = 0{\cdot}25$, N.S.; group II: $\rho = 0{\cdot}18$, N.S.); with the English letter form it is significant ($\rho = 0{\cdot}33$, $P < 0{\cdot}05$); with the Hebrew letter form it is highly significant ($\rho = 0{\cdot}44$, $P < 0{\cdot}01$). In this order the three forms of the task approach steadily closer in degree of abstraction to the matrices themselves. Interrelations between the three forms of the task would be expected and in fact the English letter form correlates significantly with the verbal form (group I: $\rho = 0{\cdot}45$, $P < 0{\cdot}05$; group II: $\rho = 0{\cdot}53$, $P < 0{\cdot}01$) and with the Hebrew letter form ($\rho = 0{\cdot}51$, $P < 0{\cdot}001$). The correlation between the verbal form and the Hebrew letter form verges on significance in group I ($\rho = 0{\cdot}40$, N.S.) but is low in group II. Table 1 shows, however, that scores on the verbal form were low in group II and restricted in range; it is perhaps surprising that so high a correlation was obtained with the English letter form in this group.

It was hypothesized above that performance on the Hebrew letter form of the task would be slower and/or less accurate than on the English letter form. The results

suggest that this hypothesis may be false, but a strict determination of the relative difficulty of the Hebrew form remains dependent on the outcome of a fully balanced experiment with six groups. Nearly every subject assimilated the Hebrew letters to vaguely similar English letters or to shapes or objects, but it remains surprising that they were able to manipulate symbols thus arbitrarily named as easily as highly overlearned English letters by the end of the experiment.

It was also hypothesized that performance on the verbal form of the task would be slower and/or less accurate than on the English letter form and this hypothesis seems to be supported by the evidence. Moreover the verbal form benefits much more from being preceded by the English letter form than vice versa. Thinking inductively in words (even the very simple words used here) was more difficult for these subjects than thinking in symbols, even unfamiliar symbols. The hypothesis attributes this greater degree of difficulty to the fact that verbal material is open to proactive interference from individual preconceptions, etc. It is also possible that more time is required to perceive verbal material and that the capacity of the short-term store for such material is smaller, but the writer knows of no directly relevant evidence that this is the case. It does not seem likely that the difference (if any) could be large enough to account for the difference in the scores obtained.

It is well known that subjects like those used in this investigation are at or near their peak as regards abstract, non-verbal intelligence and will suffer a regular decline in that capacity with increasing age. Their verbal ability is, however, likely to go on improving with age. It might be hypothesized that with increasing age, level of performance on the tasks will converge and conceivably cross over so that old people may do better on the verbal form of the task than on the English and Hebrew letter forms. The writer hopes to be able to put this hypothesis to the test.

REFERENCES

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

RAVEN, J. C. (1956). *Progressive Matrices, Sets A, B, C, D and E*. London: H. K. Lewis.

WETHERICK, N. E. (1966). The inferential basis of concept attainment. *Br. J. Psychol.* **57**, 61–9.

(*Manuscript received* 31 *October* 1966)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 227–236
Printed in Great Britain

A FREQUENCY ANALYSIS OF CONTROLLED, CONTINUOUS WORD ASSOCIATION

By W. A. MATTHEWS

Department of Psychology, University of Southampton

The responses of subjects producing continued controlled associations to eighty-seven stimuli were analysed in terms of the frequency of their occurrence as items in the language. The efficiency of hypotheses which relate word emission by subjects to frequency terms only was questioned in the light of the obtained results, and alternative proposals based on these findings are indicated.

Associative norms which have been used as the basis for experimental investigations of verbal learning, mediation and recall processes have normally been derived from single associations to the stimulus words (e.g. Russell & Jenkins, 1954; Palermo & Jenkins, 1964). The assumption has been made that the pattern of associations which are found in the norm-producing group will be reflected in the organization of the 'word store' within an individual. This 'store' is presumably involved in the processes mentioned earlier. The probability of emission of a particular associative response by an individual is assumed to be related to the frequency with which it has been produced by the norm-producing group.

Additionally, in the search for some principle which may be related to the emission of any response, the frequency of its occurrence has been emphasized. This may be assessed by either the frequency of occurrence of a word in the language as measured by the Thorndike–Lorge count (Thorndike & Lorge, 1944) or the frequency of occurrence of an item in a particular situation. These measures are usually but not inevitably correlated.

An example of this is the 'spew hypothesis' of Underwood & Schulz (1960). The spew hypothesis was formulated following the observation that the early items in a series of spontaneously emitted words are often of high frequency on the Thorndike–Lorge G count. Later words in a sequence are often of lower frequency.

It was suggested that a similar generalization is appropriate when a stimulus is presented and responses are produced to it. If this were so, the sequence of associations given by an individual should be of clearly decreasing word frequency as the order of emission increases. Single-association data are inappropriate to test this notion. However, it seems unlikely that the characteristics of a stimulus could fail to affect the responses given to it or that a verbal storage system which functions as efficiently as that of most adults could be organized predominantly in terms of frequency. The analyses to be reported here attempt to throw some light on these issues.

Method

Subjects

The subjects were students at a teacher training college (average age twenty years). There were five groups of fifty subjects, each group consisting of twenty-five men and twenty-five women.

Material and procedure

Eighty-seven stimulus words were used. They were selected with reference to a series of experiments on short-term recall in which the associative interconnexions between words composing twelve item lists were ascertained. These experiments are reported elsewhere (Matthews, 1966).

They were presented both auditorily and visually to the subjects, who were given the following instructions: 'I am going to present a series of words to you. I will say the word, and write it on the board simultaneously. As soon as you hear and see the word, I want you to write down on the sheets which you have been given as many words as you can which this word makes you think of. Only write down those words which the word with which you have been presented brings to mind. Do not associate in a sequence from one word to another from your own responses, but after each associate has been written down always go back to the initially given word.' (An example using a stimulus word not in the set was given here to show what was and was not required.) 'To help you to refer back to the given word, always write down the stimulus as soon as you have heard and seen it in the space provided on your sheet.' (The response sheets were divided into rows and columns with the column on the left reserved for the stimulus words.) 'You will be allowed 2 min to associate to each word and then you will be told to stop. You may repeat associations if you wish to do so.' (This was clarified in response to questions, but very few repetitions were in fact given.)

The responses were written on the mimeographed response sheets on which the subject had previously written the stimulus word. Each stimulus was not given to each subject. Owing to difficulties of organization five groups of subjects took part in the experiment. They were drawn from the same population and were not grouped in ways which might have influenced speed of association. Group 1 was given eight HF and seven LF stimulus words; group 2 had six HF and seven LF words; group 3 had seven HF and nine LF words; and group 4 had six HF and nine LF words. These four groups were instructed to underline the word which had just been written at the end of 1 min, and to continue associating. The experimenter notified the group by saying: 'Underline now.' Group 5 were given 27 HF and one LF words in two separate sessions.

The words used are listed in Matthews (1965, Table 1). (In this paper HF words are those occurring more than fifty times per million words sampled, except where the *AA* and *A* categories of the Thorndike–Lorge G count are specifically subdivided. LF words are always words that occur less than fifty times per million words sampled.)

Analyses

1. Each item was extracted from the protocols and a summary made of their frequency of occurrence as associations to a particular stimulus. The responses were classified in terms of their word frequency on the G count. Order of emission by the subject was ignored. A summary of the number of different associative responses, and the number of times each different response was given to each stimulus, was available after the analysis.

2. A summary was made of the first response given by each subject to each stimulus and classified in terms of their word frequency on the G count.

3. The response protocols of ten subjects (five men and five women) were randomly selected from groups 1–4 as this gave similar numbers of HF and LF words. Their associations were classified in terms of word frequency and extracted with order of emission by the subject retained.

Details from these analyses follow.

ANALYSIS 1

The mean number of different words given twice, three times, four times, etc. is shown in Table 1. Words given only once do not appear in this table. It can be seen that a greater number of different *AA* responses are given to HF than to LF stimuli, and this holds at each point noted (with a single reversal). The percentage of *AA* responses to the HF stimuli is approximately sixty. As the word frequency of the stimulus (as measured on the G count) decreases, an increasing incidence of lower-

frequency words is given. *A* responses to the same simuli show three instances of this, although only a small number of different words is produced. A greater number of different LF responses is given to LF stimuli than to HF stimuli in nine out of twelve possible instances in Table 1.

A further analysis was made to clarify the detailed occurrences at the 20+ points in Table 1, where frequent associations are summarized. The mean number of times that each different *AA* response was given to an HF stimulus was 28·4 and was 30·1 to an LF stimulus; for *A* responses, the mean to each set of stimuli was 27·6; the mean number of times that each different LF response was given to its HF stimulus was 23·4 and to its LF stimulus was 26·8.

Table 1. *Mean number of different words given to HF and LF stimuli categorized by word frequency and by the number of subjects who produced each association*

	Mean number of different words occurring as associate to stimuli								
	AA words			*A* words			LF words		
Frequency produced	HF stimuli	LF stimuli	%*	HF stimuli	LF stimuli	%*	HF stimuli	LF stimuli	%*
2	15·53	11·23	42·0	5·06	5·00	49·7	16·70	17·06	50·5
3	7·92	5·46	40·8	2·64	1·98	42·9	7·02	7·00	49·9
4	4·42	3·80	46·2	1·55	0·92	37·2	4·36	3·54	44·8
5	3·06	2·37	43·6	1·02	0·49	32·5	2·26	2·37	51·2
6	2·34	1·77	43·1	0·68	0·69	50·4	1·62	1·69	51·1
7	1·43	1·34	48·4	0·55	0·20	26·6	0·87	1·14	56·7
8	1·06	0·77	42·1	0·36	0·29	44·6	0·79	1·34	63·5
9	0·77	0·83	51·8	0·19	0·23	54·8	0·68	0·74	52·8
10	0·91	0·77	45·8	0·40	0·09	18·4	0·64	0·51	44·3
11–14	2·28	1·37	37·5	0·58	0·42	42·0	1·16	1·17	50·2
15–19	1·34	0·83	38·2	0·54	0·46	46·0	0·62	0·74	54·4
20+	2·12	1·80	45·9	0·40	0·74	64·9	0·64	1·03	61·7

* Percentage of responses made to LF stimuli.

Apparently, if LF words are given as associates they are not emitted quite so often as are *AA* words. A further detailed analysis was then made of the words given by thirty-two or more of the group of fifty subjects to the HF and LF stimuli, i.e. the extremely popular responses. Among the responses given to the HF stimuli were thirty-three different *AA* words, each produced by a mean number of 36·8 subjects; five different *A* responses given by a mean number of 37·2 subjects, and four different LF responses given by a mean number of 35·8 subjects. To the LF stimuli, nineteen different *AA* responses were produced and these were given by a mean number of 40·0 subjects, no *A* responses occurred, and seven different LF responses were given by a mean number of 36·2 subjects. This indicates that, irrespective of stimulus characteristics, the most popular responses tend to be *AA* words. This result for very frequently occurring words lends support to the spew hypothesis, if the assumption is made that words given by many subjects tend to be produced early in the sequence.

To investigate this last point, the emitted position of the 20 *AA* and 7 LF responses given very often to the LF stimuli and of the 33 *AA*, 5 *A* and 4 LF responses given very often to the HF stimuli was obtained for each subject. The means and standard

deviations were then obtained. Considering first the responses to LF stimuli, the mean emitted position of the *AA* associations was 3·92 and for the LF associations was 6·63. The S.D. of the emitted positions had been obtained for each word and the range for the twenty seven words was from 0·62 to 5·74. The lowest S.D. for an LF word was 3·63 and the highest was 5·73. Only seven scores from the twenty *AA* words were within this range. The remaining S.D.s (thirteen words) were lower than 3·63. The responses given to the HF stimuli did not fit this pattern precisely. The mean emitted position for *AA* responses ranged from 1·50 to 8·40 (S.D.s ranging from 1·34 to 6·28). For the *A* responses, the range of the emitted position was from 3·90 to 6·80 (S.D.s ranging from 3·41 to 5·30) and for LF responses the range of the emitted positions was from 2·90 to 5·50 (S.D.s ranging from 2·98 to 4·62). It is clear that responses which occur consistently in a very early position are *AA* words, whatever the frequency of the stimulus, and that these tend to have a low S.D. With both HF or LF stimuli, however, lower frequency words are often produced by many subjects but the order of production is later in the sequence and more variable. However, many responses of high frequency are also produced at points as late as or later in the sequence than the LF responses with a greater proportion of these occurring to the HF stimuli. This suggests that order of emission may depend not only on frequency characteristics but upon some other factors, e.g. strategies of responding adopted by the subject. These will be discussed in more detail in analysis 3. Nevertheless, high word frequency (as measured on the G count) appears to be a reasonably consistent characteristic of frequently produced associations. Where the responses are given by fewer subjects it is clear that simulus characteristics have an effect on the responses produced as associations, with fewer HF words being available as associative responses to LF stimuli than to HF stimuli, although even where associations occur infrequently, at least one-half of these are HF words.

ANALYSIS 2

This analysis was undertaken on the first response made by each subject to each of the eighty-seven stimulus words. As there were fifty subjects, this analysis was made on a total of 4350 words. Of these, 2526 (58 %) were *AA* words, 585 (13·5 %) were *A* words and 1239 (28·6 %) were LF words. A summary of these data, subdivided into responses to HF and LF stimuli, is given in Table 2.

The number of responses in each frequency class to the HF and LF stimuli indi-

Table 2. *Summary of the results of Analysis* 2

(a = Mean number of different initial responses by word frequency to HF and LF stimuli; b = total number of responses made to HF and LF stimuli classified by word frequency; c = mean number of repetitions per response as a function of the frequency of the stimulus and of the associative response.)

	(*a*)		(*b*)		(*c*)	
No. of responses	HF stimuli	LF stimuli	HF stimuli	LF stimuli	HF stimuli	LF stimuli
AA	12·4	7·9	29·9	27·7	2·4	3·6
A	3·3	3·3	6·6	6·8	2·0	2·2
LF	9·2	8·9	13·5	15·5	1·5	1·7

cates the predominance of very frequent words as associative responses. Stimulus frequency had no effect on the percentage of words produced which were of high or low frequency, yet the number of different words produced does show that the variety of HF associates to LF stimuli is restricted. However, Table 2*c* indicates that each of these *AA* responses to the LF stimuli was given by a greater number of individuals, when compared to those who produced *AA* responses to HF stimuli.

This finding may be compared with the point made in anaysis 1 and shown in Table 1, that fewer *AA* responses are available to LF than to HF stimuli, and suggests that one of the characteristics of LF stimuli is that these tend to elicit fewer *AA* words than do HF stimuli.

Analysis 3

This analysis was an attempt to throw more light on the spew hypothesis, which suggests that the earlier words which are emitted from a word store should be of high frequency. This notion may have been too generally applied, derived as it is from an

Table 3. *Stimulus words which produced the response sequences analysed in order of emission*

animal *AA*	kitchen *AA*	pillow 33	execution 16
article *AA*	large *AA*	joke 32	gambling 16
big *AA*	life *AA*	tailor 32	mare 14
boy *AA*	man *AA*	van 29	violin 11
cold *AA*	poor *AA*	rag 28	orchestra 10
cool *AA*	rain *AA*	suspicion 28	turnips 10
desires *AA*	show *AA*	anchor 26	warehouse 10
direction *AA*	smoke *AA*	legend 22	cranes 9
dress *AA*	stone *AA*	weave 22	crates 7
eat *AA*	Jack *A*	fabric 21	giggle 5
family *AA*	log *A*	uncertain 21	sentry 5
farm *AA*	pipe *A*	bats 19	overture 4
food *AA*	wooden 46	moors 17	buttercup 2
island *AA*	writing 39	plough 17	rugby 2
it *AA*	axe 35	trumpet 17	

analysis of Kent–Rosanoff stimuli (mainly HF words) and the first associations to these words. This may give a high incidence of HF words as first associations but a different pattern may emerge with more varied stimuli and by using controlled, continuous association. A direct comparison of the proportions of HF words as responses in analyses 1 and 2 above is inappropriate as it seems inevitable that a greater proportion of LF words must appear in continued association scores as compared to single association scores. Therefore a further count was made. It consisted of an analysis of the order of emission of the continued association responses of a randomly chosen sample of five men and five women to 59 stimuli (27 HF and 32 LF words). These are listed in Table 3, and were the words used as stimuli for groups 1–4 inclusive (see Method).

An extreme version of the spew hypothesis would suggest that each individual should first produce HF responses followed by LF responses. A Wald–Wolfowitz runs test (Siegel, 1956) was carried out on the 590 sequences produced by the ten subjects to the fifty-nine words with HF (*AA* and *A*) and LF words as the two categories. Only eighteen from 590 tests reached the 0·05 significance level and this

indicates that the order of emission was not determined by the frequency of the responses. A second analysis was undertaken. The first emitted response of the ten subjects to each stimulus was noted and classed as either high (*AA* or *A*) or low frequency and a binomial test was then applied to these first responses of the ten subjects to ascertain whether the number of HF words occurring in the ten words was likely to have been obtained by chance. A similar procedure was applied to the ten second responses, ten third responses and so on until the associations of at least one subject were exhausted. Artifacts due to the varied length of each subject's associative sequences must have been present later in the series but the results for the first six words produced by each subject appear in Table 4.

Table 4. *Number of sequences of first, second and nth responses to each stimulus word given by ten subjects which contained a greater nunber of HF words than would be expected by chance*

(Total number of possible sequences for each column for the HF stimuli is 27 and for the LF stimuli is 32.)

		Order in which word was emitted by subject					
	Significant at	1	2	3	4	5	6
HF stimuli	0·001 level (all HF words)	2	2	—	1	—	—
	0·01 level (9 from 10 words)	7	5	1	—	1	1
	0·05 level (8 from 10 words)	5	5	6	3	4	2
LF stimuli	0·001 level (all HF words)	1	—	—	—	—	—
	0·01 level (9 from 10 words)	4	1	—	1	1	—
	0·05 level (8 from 10 words)	13	6	5	4	4	1

A rapid decline in the number of sequences of eight, nine or ten HF responses occurs and this is not related to any particular stimulus class except for the second word produced by the ten subjects to the LF stimuli. As is clear from Table 4, there is a sudden drop in the percentage of HF responses produced at this point. This lends support to the findings from analysis 1, which used all the responses from the fifty subjects to all the stimuli, that many HF words are given as responses to LF stimuli, that these may occur early in the series, but that additionally a greater number of LF words are produced to LF than to HF stimuli.

The details of this analysis emphasized the importance of the semantic characteristics of both stimuli and responses (cf. Matthews, 1965). The pattern which was observed in the response sequences of many subjects suggests that the basic processes may consist of the production of a subset of semantically related words with other subsets, which may or may not be partially overlapping, appearing in the sequence. Word frequency may be a subsidiary factor within each subset. For example, to the stimulus 'farm', early responses may range over animals, and be followed by produce, buildings, workers and so on. The 'runs' tests and the results from the binomial tests applied to the first few associations which had been produced could be interpreted as giving some support to this hypothesis. It is of interest that Pollio (1964) reported temporal grouping in the emission of responses by his subjects in a free-association situation and this suggests that temporal boundaries may correlate with

semantically related subsets. Pollio does not analyse semantically, and the group method by which these written associations were obtained prevented temporal measurements being made. However, attempts to produce precise classificatory labels for each subset were unsuccessful, there being frequent ambiguities about the length of each subset, and the semantic category to which the items in the sequence could be related.

An attempt was made to obtain some empirical support for this proposal and this is reported in analysis 4.

Analysis 4

Method

Sixteen stimulus words and the associations which had been produced to them were selected. Eight of these words were of high (*AA* or *A*) and eight of low word frequency. Four of the words had been presented to group 1, and four each to groups 2, 3 and 4 respectively (see analysis 1). The associative responses of four individuals to each stimulus were selected at random. Thus sixty-four sets of associations were extracted, four sets to each stimulus word, and with a single individual contributing responses to four different stimuli. Each sequence was typed in the order of emission on a separate strip of paper. Two of the four sets of the associative responses to each stimulus were selected and the thirty-two sequences which resulted were arranged randomly and stapled together to form a booklet. The remaining thirty-two sequences were arranged into a separate booklet. Control lists were constructed in the following way. An analysis of the number of words in each of the sixty-four associative sequences was made. The range was from five to forty words. The number of sequences occurring in each five-word group between zero and forty words was obtained and the mean number of HF and LF words of nouns, verbs and adjectives in these associative sequences obtained. Control lists were selected at lengths 5, 9, 15, 17, 24, 27, 32 and 37 words, these being the mean length of the sequences within each five-word subgroup. The proportion of HF and LF words within the three grammatical classes which had been obtained by the method described above were included in each sequence. The words were selected from the Thorndike–Lorge G count, by a random choice of page, and of position on the page, and by a continual perusal down the column until a word which met the above criteria appeared. Two lists of each length were constructed, consisting of different words. Each list was typed in two different arrangements. Sets of these sixteen lists were arranged in a random sequence and stapled together to form a booklet. For the experimental lists (two different sets of thirty-two sequences) and the control lists (one set of 16 sequences), ten copies of each booklet were made and given to different subjects. These were graduate or undergraduate members of the University of Southampton. Accompanying each booklet were the following instructions:

'The following sets of word sequences are presented to you to enable you to make judgements on them in terms of their conceptual reference. (This will be specified more precisely later.) Will you inspect each sequence, then write a number, beginning with 1, over any words in the series which appear to you to be closely conceptually related or have similar reference. For words with different reference use different numbers (2, 3, 4, etc.). Words which are classified as belonging to the same category may appear consecutively or not in the sequence. If none of the words are conceptually related, write a different number over each word.'

'Conceptually related' is a vague term but may be defined as (1) 'having attributes common to a class' and (2) 'referring to the same group of ideas, or objects or events or processes'.

Three examples were given and one of these follows:

1	2	3	3	4	4
field	farmer	plough	tractor	dinner	food
4	1	4	4	4	4
eat	meadows	plate	dinner	pie	gravy

No time-limit was imposed.

RESULTS

The sixty-four associated sequences, each of which was classified by ten subjects, contained a total of 10,430 words (in 640 lists). From these, 6008 were judged to be grouped into sequences of two or more words, a further 1024 isolated words were labelled with the same category numbers which had occurred elsewhere in a sequence; the remaining 3398 words were judged to be not conceptually similar to words occurring in the lists in which they occurred. The large number of words which were judged to be related to categories and the high proportion of these which occurred in sequences of two or more words give general support to the hypothesis being proposed. However, the total number of words grouped gives no information about the patterns which occurred within any particular associative string, and a detailed inspection of the sequences showed that there were frequent recurrences of categories after differently categorized sequences had intervened. Thus 862 (14·3 %) of the 6008 categorized words which were in sequences occurred in this way, and a further 900 words (15·0 %) were categorized as occurring in a sequence after the occurrence of an isolated word which had been classified in an identical way. There were thus 4246 words (70·6 % of the total number of the categorized words which appeared in a sequence) which were part of categorized sequences which were not repeated within the set of associations given by a subject. The mean length of each categorized sequence which occurred without a prior sequence and without a succeeding one was 3·02 words; where a sequence occurred after an isolated similarly categorized word, the mean length was 2·96 words, and where it occurred after an earlier sequence was 2·76 words. A sign test (Siegel, 1956) was used to assess the significance of any differences in direction which occurred between the mean number of words categorized in an identical way where one sequence only occurred (3·03 words) and where a sequence followed an earlier one in the same list which had been categorized in the same way (2·76 words). These differences were significant ($P < 0{\cdot}02$). The remaining two possible comparisons did not show significant differences.

The control lists served two purposes. Firstly, to investigate the percentage of words in a randomly chosen sequence which could be classified as belonging to a particular conceptual category, and secondly to observe what proportion of these similarly categorized words occurred in a sequence. The highest percentage of words placed in categories in any control sequence was 42·5 %. This was greater than only nine from the sixty-four lists produced by continued association. (The differences were significant on a Mann–Whitney U-test, $P < 0{\cdot}001$.) Additionally, a comparison was made between the number of categorized words which were grouped in the control and associated lists. This was done using fourteen control (two control lists scored zero on the first measure) and nine associated lists which did not differ significantly in the percentage of categorized words in each (Mann–Whitney U-test, not significant). Thus the associated lists which contained the smaller proportion of categorized words were used. There was nevertheless a significant difference (Mann–Whitney U-test, $P < 0{\cdot}001$) between the percentage of categorized words appearing in a sequence, with more words reported sequentially in the associated than in the control lists.

There is then general support for the hypothesis that associated word sequences may be reasonably classified into semantic categories and that these tend to appear

in sequences. Sequences which are similarly categorized may be repeated within the same set of associations, but where there are no repetitions the mean length of a sequence of similarly categorized words is reliably greater than where a sequence occurs after an earlier similarly categorized single word or sequence. One additional point of interest is that the responses to the adjectives (big, large, poor, wooden) within this set clearly produced a smaller proportion of grouped responses than did those made to stimuli which could be classified as nouns, verbs or both.

Discussion

Some support has been evinced from these analyses for the proposal that early responses to verbal stimuli tend to be HF words and this applies whatever the word frequency of the stimulus. However, this generalization does not account for the frequent occurrence of HF words (within a particular sequence) after the occurrence of LF words (cf. analysis 3). It was proposed that some principle of semantic reference may be involved in the production of words in this way, and that any change in this semantic reference may also lead to a recurrence of HF words when the new sequence begins. Analysis 4 gives some support to this notion. Here groups of words were classified semantically and several different classifications within any one sequence often occurred. Clearly a response to a stimulus may be initially of high frequency, with semantically related LF words following. Further sequences which are classified differently follow. Sixty-three per cent of the semantically related sequences which occurred on the sheets classified in analysis 4 began with an HF word. Fifty-two per cent of the sequences which were classified as repeats of a categorized string that had already occurred within a particular associative sequence began with HF words. Where nouns are the stimuli, the associations seem to be grouped more easily than where the stimuli are adjectives or (to a lesser extent) verbs. This is understandable with this sample of stimuli, for these adjectives (big, large, poor, wooden) may occur with a wide variety of different words, and a larger proportion of these than of the situations in which the noun stimuli occur may be unrelated. If this were the case, we would expect to find a greater proportion of the responses to adjectives than to nouns or verbs, consisting of single, non-conceptually related words. Some support is found for this proposal in that forty-three from the forty-eight noun or verb stimuli have less than one-half of the associates classified as unrelated to any other words in the same associative sequence compared with ten from the sixteen associative sequences to adjectival stimuli.

In continued controlled association, the following system may operate. The first response to the stimulus is determined by the frequency with which a particular item has earlier been produced as an overt or covert response to it. Subsequent associations are produced with the subject monitoring both the initial stimulus and the subsequently produced responses. This may be analogous to the situation manipulated by Howes & Osgood (1954) in which specific associations became increasingly probable as the number of preceding stimuli was increased and their semantic reference was limited. As the number of responses given to the stimulus increases, the probability of the emission of LF words should increase and the semantic reference should narrow until all the words stored within this limited part of some hypothetical

word storage system have been emitted. A change of semantic reference should follow, with the frequency of the initial words again being high as a different semantically classifiable series is produced A further prediction may be made. The occurrence of this semantically classified sequence in the verbal (or otherwise) experience of the subject should be less than his experience of the initially produced sequence, but the control over experience necessary to isolate these factors will be extremely difficult, if not impossible, to achieve. This may be modified where a sequence of associations with an initial HF response occurs but without a decline in the word frequency of the associations in the series. The probability of switching to quite different semantically classifiable items should be higher than where there is a decline in the word frequency of the items in the sequence (assuming that many HF words have a greater number of different associative connexions than do LF words). It may also be expected that there would be greater probability of the recurrence of an already produced semantically related sequence in this type of case than where there had been a decline in word frequency over a sequence. There was no clear empirical support for this proposal from data.

What seems to be clear from these analyses is that any attempt to describe the characteristics of emitted associative sequences in frequency terms will be inadequate. In order to systematize these data, referral to some categorization process was necessary, and even within these categorical subgroups, a single word frequency approach did not account for the incidence of category groups which were repeats of earlier sequences, within the same associative string. Nor does a frequency approach commend itself to explain the ordering of particular semantic sequences within an associative string. Proposals to account for the latter are put forward but the control over experience necessary to isolate these factors will be extremely difficult, if not impossible, to achieve.

REFERENCES

HOWES, D. & OSGOOD, C. E. (1954). On the combination of associative probabilities in linguistic contexts. *Am. J. Psychol.* **67**, 241–58.

MATTHEWS, W. A. (1965). The relation between association norms and word frequency. *Br. J. Psychol.* **56**, 391–9.

MATTHEWS, W. A. (1966). Continued word associations and free recall. *Q. J. exp. Psychol.* **18**, 31–8.

PALERMO, D. S. & JENKINS, J. J. (1964). *Word Association Norms.* Minneapolis: University of Minnesota Press.

POLLIO, H. H. (1964). Composition of associative clusters. *J. exp. Psychol.* **67**, 199–208.

RUSSELL, W. A. & JENKINS, J. J. (1954). The complete Minnesota norms for responses to 100 words from the Kent–Rosanoff association test. (Technical Report no. 11, University of Minnesota, Contract N 8-onr-66216.)

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words.* New York: Columbia University Press.

UNDERWOOD, B. J. & SCHULZ, R. W. (1960). *Meaningfulness and Verbal Learning.* Philadelphia: Lippincott.

(*Manuscript received* 8 *November* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 237–242
Printed in Great Britain

PHRASE UNITS AND THE RECALL OF GRAMMATICALLY STRUCTURED NONSENSE

By P. HERRIOT

Department of Education, University of Manchester

Inflexions and function words were added to nonsense stems to form strings fitting three different sentence frames. These strings were then divided into phrase units, and also into units of the same size but with the inter-unit boundaries not occurring at the phrase boundaries. Other conditions introduced the randomization of the order of the units, and also of the nonsense stems within units. Results, in terms of the number of trials to criterion, showed that there was a significant difference between the two groups learning the strings in the correct order, one in phrase units and the other in non-phrase units; there was also a significant order effect, the strings with stems randomized being harder to learn than the strings with units randomized and the ordered strings. When transitional error probabilities were calculated, it was found, contrary to expectation, that error probabilities were higher between the non-phrase units than between the phrase units. It was concluded that phrase units are basic in the recall of sentences, and that the sequential organization of stems is most important within phrase units.

Epstein (1961) found that strings of nonsense stems with inflexions and function words added were learned more rapidly when arranged in sentence frames than when arranged in random order. He further found that the presence of inflexions led to quicker learning than their absence. He concluded that grammatical structure aided recall, and that the amount of structure present was a relevant variable. However, the nature of the structure could not be inferred; randomization of the order of the words could be destroying the left-to-right sequential dependencies between words, or it could be disrupting higher order units and two-way dependencies. Epstein (1962) further found that no such significant differences occurred when the individual items of the strings were presented in sequence one by one on a memory drum. He concluded that the facilitatory effect of grammatical structure on recall was due to some sort of structural, as opposed to linear, analysis of the strings by subjects. However, the delay involved in the method of presentation may have affected this result, and the subjects may not have realized that the items were in any way connected. If one accepts the methodology, it must be concluded that relationships other than left-to-right dependencies were involved. However, the importance of left-to-right dependencies was shown by Epstein's (1963) finding that grammatically structured material was better learned than randomized material when left-to-right learning was explicitly required; but that randomized material was better learned than structured material when right-to-left learning was required. However, it is clear that the randomization of stem order disrupts any larger structural units which may exist in addition to stem dependencies. It therefore cannot be said that Epstein has made clear the nature of these units or their interaction with stem dependencies.

Saporta (1965) found that sequences derived from five letters of the alphabet were learned with fewer errors when arranged in constituent units resembling phrases than in randomized units or with equal spacing between each letter. Johnson (1965) reported that, in the learning of simple English sentences, most errors in recall

occurred at the transitions between phrase units, and that the larger the constituent unit which followed a transition, the more likely an error was to occur. Johnson (1966) further found that when an adjective and a noun were learned by a paired-associate method, fewer transitional errors occurred between them when they were recalled as part of a sentence than if there had been no such previous learning. However, the same did not apply to a noun–verb pair. The conclusion was that the improvement did not occur in the latter case because the noun and the verb were separated by a major constituent boundary, while the adjective and noun were not. These findings support the hypothesis that subjects use phrases as basic units in the memorizing and recall of sentences.

It was therefore hypothesized that Epstein's results were due to the possibility of dividing the strings of nonsense stems into phrases, once grammatical structure was supplied by the added function words and inflexions. His findings regarding the importance of sequential stem dependencies suggest that intra-phrase stem dependencies are important.

An experiment was therefore set up in which strings were generated and then divided into units. These were phrase units, or units of the same size but with the breaks occurring at points other than the constituent boundaries. It was hypothesized that the former would require fewer trials to a criterion of perfect recall than the latter. A randomization procedure applied to units was expected to lead to better recall than the randomization of stems within units, thus indicating the importance of sequential dependencies of stems within units. It was further predicted that transitional error probabilities (TEPs) would be greater at phrase-unit boundaries than at non-phrase-unit boundaries.

METHOD

Subjects

The subjects were seventy-two graduates attending courses in education. None possessed degrees in psychology, nor had any attended courses on the psychology of language. The subjects were randomly assigned to six groups of twelve.

Design

Of the six groups, three learned strings divided into phrases (P) and three learned strings divided into other units (NP). Within these two sets of three groups, the first group learned the strings in the correct order throughout (O), the second learned with the order of the units randomized (UR), and the third with the order of the stems within the units randomized (SR).

Material

The material was derived from the bigram frequency table in Underwood & Schulz (1960) in the manner described by Forster (1966). Twelve strings, of seven different nonsense stems each, were generated, e.g. *chim teng elan gemor lev tirnal ralig*. The order of the seven stems was randomized eleven times, so that there were in all twelve versions of each of the twelve strings. Three strings were then selected at random for each group, each subject within a group receiving a different version of the same three strings. Each one of these strings was then fitted to one of the following three sentence frames, which are those used by Forster.

(*a*) The ———y ———s ———ed the —— ————ly with ——— and ———s.
(*b*) The ———s who ———ed the ———s ———ly ———ed the ——— ———.
(*c*) The ——— ———ed to the ———'s ———est ——— who ———ed ———ing.

An example of a version of a string fitted to a frame is: *The chimy tengs elaned the gemor levly with tirnal and raligs.*

These versions of strings fitted to sentence frames (referred to hereafter as 'strings') were typed horizontally in capital letters without punctuation marks (apart from an apostrophe) on strips of cardboard measuring $8 \times \frac{1}{2}$ in. The division into units was carried out by means of sets of brackets enclosing asterisks. Examples of strings presented to each group are:

PO: THE CHIMY TENGS (**********) ELANED THE GEMOR LEVLY (**********) WITH TIRNAL AND RALIGS

NPO: THE IZEFS WHO (*********) TROKED THE (********) TIDUNS POSELY IRSTED THE (********) STIL FEADOIL

PUR: GODED THE AST JOPLY (**********) WITH SEVY AND YONKARS (*******) THE ICANCY EHALS

NPUR: VIPED DASCING (***********) JUDIL'S DIXEST PEHIMO WHO (********) THE YEV GACED TO THE

PSR: JUKARS THE (*******) THE WHO EGOES WEXED (********) FUGED MUPELY (*********) ONUCOB EW THE

NPSR: RAITED THE JOLUR'S THE TO BINK (*********) YOSEB WHO MOBEMEST (***********) HOBONING OBINED

Procedure

The experiment was administered in groups ranging in size from four to twenty-five subjects. The subjects were instructed not to look at their three cards until told to do so. They were told that the object of the experiment was to test their memory for strings of nonsense stems; that they were to aim at perfect reproduction of order and spelling; and that when they were completely certain that they had correctly recalled a string, they were to check their version with the card on the next trial. If correct, they were to proceed to the next string. They were instructed to learn the strings in the units into which they had been divided by the experimenter.

On instructions, the subjects turned over any one of the three cards and read it for 8 sec. They then turned the card back and wrote as perfect a recall as they could in 30 sec. During the 10 sec period which followed, the subjects folded back the attempt they had just completed so that it was not visible on the next trial. The process was repeated until the subjects had completed the three strings. The criterion was perfect recall, but it was found that, despite instructions, the final attempts of many subjects still contained errors. Where there were three or less stems containing errors, one trial was added for each stem in error. In the case of four subjects, more than three stems contained errors and their results were discarded.

Results

Results were calculated in terms of number of trials to a criterion of recall. The total number of trials for each subject was calculated rather than the mean number per string. This was because Martin & Roberts (1966) showed that sentences of greater Yngve depth (defined as the mean number per word of the left branches of a binary phrase-marker tree) were harder to recall. Sentence frame (*c*) above is of greater depth than (*a*) and (*b*), and on analysis suggestive but non-significant differences were found. Table 1 therefore shows the mean number of trials to criterion for the six groups for all three strings combined.

Two-way analysis of variance was not applied, since it was not hypothesized that there would be significant differences between phrase and non-phrase units when the order of units and of stems was randomized. A *t*-test was performed on the PO and NPO groups ($t = 2{\cdot}41$; $P < 0{\cdot}05$). A further eighteen subjects were added to each of these groups to confirm this result. These additional subjects were given different strings from those of the original subjects. When the two resulting groups of thirty subjects were compared, the results were again significant ($t = 2{\cdot}91$; $P < 0{\cdot}01$). One-way analysis of variance was performed on the PO, PUR and PSR

groups. Significant variance was found ($F = 13{\cdot}48$; $P < 0{\cdot}005$) and t-tests were therefore carried out. The PO group differed significantly from the PSR group ($t = 4{\cdot}93$; $P < 0{\cdot}001$); so did the PUR group ($t = 3{\cdot}87$; $P < 0{\cdot}001$). For NPO, NPUR and NPSR groups, $F < 1$.

Table 1. *Mean trials to criterion*

		O	UR	SR
P	Mean	12·6	14·2	19·9
	S.D.	3·17	3·51	3·75
NP	Mean	16·0	15·4	18·0
	S.D.	3·76	4·82	5·57

Table 2. *Mean TEPs for two groups*

	Transition			
	3	4	7	8
PO	0·178	0·014	0·124	0·298
NPO	0·135	0·183	0·029	0·533

TEPs were then calculated for the transitions between units for groups PO and NPO. A TEP is defined as the frequency that the word following a transition is incorrect, given that the word preceding it is correct, divided by the total frequency with which this preceding word is correct. There were seven transitions for each subject. In the case of the PO group, these occurred, of course, at the junction between phrase units, whereas in the case of the NPO group they occurred at junctions between units of smaller constituent size. It was therefore predicted that there would be a higher TEP in the case of the PO group, in accordance with Johnson's findings, but this was not so. The mean TEP for each group for all seven transitions was: PO: 0·217; NPO: 0·344. The Mann–Whitney U-test was performed and a significant difference was found on a two-tailed measure ($U = 28$; $P < 0{\cdot}02$). It was immediately perceived that these results might be due to the greater number of nonsense stems which followed the transitions in the case of the NPO group. For this group six nonsense stems and one function word followed the seven transitions, whereas for the PO group the ratio was 3:4. The type (a) strings were therefore analysed on their own, since in both groups one function word and one nonsense stem followed the two inter-unit transitions they contained. Table 2 shows the mean TEPs for four transitions; these are the transitions at which an inter-unit break occurs in either of the two strings. Mann–Whitney U-tests give significant differences between groups for transitions 4 and 8 ($U = 40, 36$; $P < 0{\cdot}05$, $P < 0{\cdot}025$ respectively), but just fail to show significance for transitions 3 and 7 ($U = 45, 48$). This difference cannot be due to function words rather than nonsense stems following the transition, since transitions 3 and 8 are followed by nonsense stems, 4 and 7 by function words. It shows that the expected higher TEP for group PO at phrase-unit junctions does not occur so much as the reverse: TEPs are significantly higher only for the NPO group at non-phrase junctions.

Discussion

The significant difference in terms of trials to criterion between the PO and the NPO groups indicates the importance of phrase units in the learning of any material which has grammatical structure. Because of the different experimental procedure involved, it was not possible to compare the PO group with a group learning the strings undivided in any way. However, it may be hypothesized that there would be only marginal differences between the two, caused by the printed divisions enabling the PO group to perceive the structure more quickly. It is also clear that the significant differences between the P and NP groups occurred only when the strings were ordered and not when randomization of units or stems had taken place. It is tempting to infer that the phrase units only assist in the organization of the material when the material is itself in the form of a sentence, although obviously such inference from a null finding is dangerous.

The importance of the order of stems within units may be inferred from the significant difference between the PO and PSR groups. Epstein's procedure had randomized stems within strings, not stems within units. This intra-unit order effect is further emphasized by the finding that there was a significant difference between the PUR and PSR groups; this result shows that the difference between the PO and PSR groups is not due to randomization in general but to the randomization of stems within units in particular.

An explanation of the TEP findings is more difficult. Given Johnson's results, it must be hypothesized that these apparently contradictory findings are due either to the use of nonsense stems in the present experiment or to the experimental manipulation of units. The latter involves a visible breaking up of the string in ways either in accordance with (PO) or in opposition to (NPO) subjects' habitual ways of organizing sentences. When the subjects learned in non-phrase units, the markers* on which they habitually depended for recognition and memorization of units were not in their usual position. To take an example, the NPO group's type (*a*) string was THE OMEPY STERIES NAUTED (**********) THE JIZE SELIDLY WITH (********) SIPIV AND TUVOES; in this case the words NAUTED and WITH are taken from their places at the commencement of a phrase unit and placed at the end of the previous unit. NAUTED should act as a marker at the beginning of a predicate, indicating that an object and perhaps an adverb are to follow. Similarly, WITH should act as a marker at the start of the adverbial phrase, indicating the nature of the phrase and the fact that a noun is to follow. When these markers are removed, it is probable that the rest of the unit for which they should be functioning as markers will not be so well remembered. If one assumes that part of the function of markers is to act as a cue to memory and as a clue to expectations, then it is reasonable to suppose that when they are removed, the unit and, in particular, the next item in the unit would be far harder to remember. This would explain the higher TEPs in the case of the NPO group, since the markers have been removed from the front of their phrase units. If this is the explanation, investigation is required into the nature and function of markers.

* The term markers is used to refer to certain inflexions and function words; it does not refer here to the phrase-structure tree itself, nor to the corresponding set of rewrite rules.

REFERENCES

EPSTEIN, W. (1961). The influence of syntactical structure on learning. *Am. J. Psychol.* **74**, 80–5.

EPSTEIN, W. (1962). A further study of the effect of syntactical structure on learning. *Am. J. Psychol.* **75**, 121–6.

EPSTEIN, W. (1963). Temporal schemata in syntactically structured material. *J. gen. Psychol.* **68**, 157–64.

FORSTER, K. I. (1966). The effect of syntactic structure on non-ordered recall. *J. verb. Learn. verb. Behav.* **5**, 292–7.

JOHNSON, N. F. (1965). Linguistic models and functional units of language behaviour. In *Directions in Psycholinguistics*, ed. S. Rosenberg. New York: Macmillan.

JOHNSON, N. F. (1966). The influence of associations between elements of structured verbal responses. *J. verb. Learn. verb. Behav.* **5**, 369–74.

MARTIN, E. & ROBERTS, K. H. (1966). Grammatical factors in sentence retention. *J. verb. Learn. verb. Behav.* **5**, 211–18.

SAPORTA, S. (1965). Grammatical models and language learning. In *Directions in Psycholinguistics*, ed. S. Rosenberg. New York: Macmillan.

UNDERWOOD, B. J. & SCHULZ, R. W. (1960). *Meaningfulness and Verbal Learning*. Philadelphia: Lippincott.

(*Manuscript received* 2 *December* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 243–245
Printed in Great Britain

ORDER OF RETRIEVAL OF INFORMATION SIMULTANEOUSLY PRESENTED TO EAR AND FINGER

By MILLARD C. MADSEN and LANNY SIMS
University of California, Los Angeles

Pairs of letters of the alphabet were simultaneously presented to the ears and fingers of eight blind Braille readers. Order of retrieval following 0·5, 1·0 and 2·0 sec interpair intervals proceeded by modality of input. There was a greater, but not statistically significant, tendency for subjects to retrieve by order of input after a 4 sec interpair interval. The results are consistent with Broadbent's (1958) theory, but indicate a greater tendency to retrieve by modality after longer interpair intervals than has previously been indicated for other sensory combinations.

In the study of memory storage processes, a number of experiments have been reported in which the method of simultaneous stimulation of two sense organs was used. The original study by Broadbent (1954) involved simultaneous input to two ears; subsequent studies have involved simultaneous input to the two eyes (Sampson & Spong, 1961) and to eye and ear (Broadbent & Gregory, 1961; 1965). The purpose of the present study was to examine the relationship between presentation rate and order of retrieval with a sensory combination not previously studied in this context, auditory and tactile. The problem of presenting equivalent information through auditory and tactile channels was overcome by using blind subjects who were competent Braille readers. Thus a letter of the Braille alphabet could be presented to the finger while the experimenter simultaneously pronounced a different letter. The interval between simultaneously presented pairs could then be varied in a manner comparable to the procedure used in experiments involving other sensory combinations.

The results of Broadbent's (1954) original study indicated that upon retrieval subjects tended to group items by ear when the interval between successive pairs was 0·5 or 1 sec, but at 1·5 and 2 sec intervals there was an increasing tendency to retrieve items in the actual order of arrival. It was on this basis that subjects in the present experiment were tested at interpair intervals of 0·5, 1 and 2 sec. When the expected results did not appear, subjects were retested at a 4 sec interpair interval.

Method

Eight blind subjects, five male and three female, who were students and staff members at the Braille Institute in Los Angeles, participated in the experiment. Ages ranged from nineteen to sixty-eight years and length of blindness ranged from four months to fifty-four years. Two subjects were congenitally blind. No diabetics were included in the study, since their tactile sensitivity is severely diminished. A Tellatouch Braille typewriter (American Foundation for the Blind) was used to present Braille letters to the finger. Auditory stimuli were pronounced by the experimenter. Each subject was familiarized with the Tellatouch machine and given practice trials on the recognition of Braille characters. Only those who could consistently identify Braille characters when presented for 0·5 sec were included in the experiment. This requirement proved to be a rigid criterion, as a number of available subjects were eliminated from the study.

Subjects were told that they would receive six letters of the alphabet, three in Braille and three aurally, and that they were to repeat as many of the six letters as possible in any order. A practice trial was given to ensure that the task was understood. Subjects were initially given thirty trials. These consisted of ten trials each at interpair intervals of 0·5, 1 and 2 sec. The stimulus sets of six letters were selected randomly from the alphabet without replacement. The order of presentation of letter sets and the order of interpair intervals were independently randomized for each subject. All subjects were given ten trials with a 4 sec interpair interval approximately 2 weeks after the initial testing session.

RESULTS

Responses were scored as paired, modal or mixed. Responses were scored as paired when order of retrieval proceeded in the order of original presentation, as modal when all of the letters presented to one modality were retrieved before any of the letters presented to the other modality, and as mixed when order of retrieval indicated elements of both modal and paired responding. If, for example, the ears received ANF and the finger received EPD, a mixed order of retrieval might be ANEPDF. The order of retrieval was inferred for erroneous and incomplete responses. If, with the above input, the order of retrieval were ABNOFS, the response was scored as paired. If the order of retrieval were ANFBOS, the response was scored as modal. The subject's strategy is quite clear in both cases despite the fact that three errors were made.

Table 1. *Mean number of letters recalled correctly by modality of input (upper) and mean responses by each order of retrieval (lower)*

	Interpair interval (sec)			
	0·5	1·0	2·0	4·0
Auditory	2·65	2·79	2·70	2·88
Tactile	1·66	1·79	1·94	1·76
Modal	9·62	8·37	9·12	6·50
Paired	0·13	0·63	0·63	2·13
Mixed	0·25	1·00	0·25	1·37

The mean number of letters recalled correctly for each modality and the mean number of responses by each retrieval order at each interpair interval are presented in Table 1. It is obvious that there were many more errors for tactile than for auditory stimuli. This was true for all subjects at all intervals. An auditory letter was retrieved first on 76 per cent of all trials and a Braille letter was retrieved first on 24 per cent of all trials. Seven of the eight subjects reported more auditory than tactile letters first.

Because of the small number of responses in which retrieval was by any order other than by modality, paired and mixed responses were combined for statistical analysis. An analysis of variance of mean responses in which any switching took place indicated no significant differences between means attributable to interpair interval ($F = 3{\cdot}29$; D.F. 3, 28; $F = 3{\cdot}34$ at $P = 0{\cdot}05$). No further statistical tests were performed. The important result, however, is not that there was a tendency toward fewer responses by modality at the 4 sec interval, but that there were so few responses in any order other than by modality at any interval.

Discussion

These results extend the original phenomenon reported by Broadbent (1954) to another combination of sensory input. However, the results are not in strict accord with what would be predicted on the basis of Broadbent's (1958) theory, which holds that digits are stored by sensory channel at rapid presentation rates because attention cannot be switched rapidly enough to process the digits in serial order. This theory accounts for the present results at the more rapid presentation rates. Since the majority of responses at the 4 sec rate were also by modality, this indicates that switching of attention from channel to channel does not necessarily occur even when there is sufficient time to switch.

Yntema & Trask (1963) demonstrated that switching of attention at rapid rates of presentation does occur if there is some reason for it. In their experiment, items could be categorized not only by the source of arrival, but also by the meaningfulness of the stimuli. Their subjects were more adept at retrieving by type of material (words *vs.* digits) than by source or pair, in spite of the fact that subjects had to cross channels in order to preserve meaning. Thus it was hypothesized that the items were more reliably tagged as words or digits than by source of arrival. It also would appear likely that when source of arrival is the only characteristic which tags the incoming items, then the relative strength of the tag will vary with different sensory combinations. The literature on simultaneous stimulation indicates a possible hierarchy of such tag strength. Sampson & Spong (1961) found that response by pair is typical when digits arrive simultaneously to the two eyes. Response by source is more frequent when the two ears are used (Broadbent, 1954) and still more frequent when the sources are eye and ear (Broadbent & Gregory, 1961, 1965). The present results indicate that information is even more strongly tagged by source when arrival is by ear and finger. Thus subjects in free recall retrieved the letters by modality, even when sufficient time to switch attention and store by serial order was available.

The authors are indebted to Mr Howard F. Bergherm of the Braille Institute of America, Inc., Los Angeles, California, for providing facilities and subjects for the experiment.

References

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *J. exp. Psychol.* **47**, 191–6.

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon.

Broadbent, D. E. & Gregory, M. (1961). On the recall of stimuli presented alternately to two sense-organs. *Q. J. exp. Psychol.* **13**, 103–9.

Broadbent, D. E. & Gregory, M. (1965). Some confirmatory results on age differences in memory for simultaneous stimulation. *Br. J. Psychol.* **56**, 77–80.

Sampson, H. & Spong, P. (1961). Handedness, eye-dominance, and immediate memory. *Q. J. exp. Psychol.* **13**, 173–80.

Yntema, D. & Trask, F. P. (1963). Recall as a search process. *J. verb. Learn. verb. Behav.* **2**, 65–74.

(*Manuscript received* 6 *December* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 247–252
Printed in Great Britain

VISUAL IMAGERY AND THE ORGANIZATIONAL PROPERTIES OF PERCEIVED STIMULI

By PETER W. SHEEHAN*
University of Sydney, Australia

Seventy-two subjects constructed blocks to indicate their perceptions and imagery of geometrical patterns which varied in their meaningfulness and stimulus complexity. It was hypothesized that vivid and poor imagers process visual information differently and that their styles of perception lead to differences in retentive behaviour. Results supported the hypothesis that vivid imagers perceive literally while poor imagers use coding devices to organize their perceptions.

A recent series of studies by the author (Sheehan, 1966) has shown that vivid and poor imagery reflect different degrees of correspondence between imaging and perceptual behaviour. Vivid imagers accurately reproduce stimuli as they perceive them but poor imagers are inaccurate in their imagery. There is considerable evidence that encoding processes are important sources of variation in perception and memory (Averbach & Sperling, 1961; Haber, 1964; Mackworth, 1963). Subjects using different coding strategies differ in the accuracy of their stimulus descriptions in recall (Haber, 1964). Such evidence suggests that the differences in the retentive behaviour of good and poor imagers are related to individual differences in the organizing of visual information at perception.

The work of Gardner (1953) and Bruner, Goodnow & Austin (1956) has long drawn attention to the role of individual differences in the coding of visual information. Subjects may organize data at perception by using coding or mnemonic devices or may perceive in a photographic or literal way. If organizers and literal perceivers are different types of subjects, then recall with vivid imagery may emphasize details taken by literal perception, whereas recall with poor imagery may emphasize details retained by the use of coding or organizing devices. According to this hypothesis poor imagers who attempt to code symbolically the original stimulus perceive details, but it is the coding device which is retained and details of the stimulus are lost at encoding and decoding. On the other hand, vivid imagers who literally perceive the stimuli element by element are better able to account for the complexity and record the stimulus features accurately in their imagery.

The present investigation was concerned with the effect on imaging behaviour of varying the organizational properties of perceived stimuli. It made predictions from the assertion that vivid imagers perceive literally while poor imagers code for retention. The investigation employed two stimulus displays which varied in their degree of patterning or amenability to coding; and subjects known to be vivid and poor imagers were asked to reconstruct them. There were three conditions (1, 2, and 3) of each stimulus display. Colour of stimulus elements was varied in condition 1; colour and shape of elements were varied in condition 2; and colour, shape and size of elements were varied in condition 3. Variation in attributes was designed to increase

* Now at The City College, The City University of New York.

stimulus complexity. It is known that coding is more difficult when stimulus complexity is increased by interdimensional variability (Battig & Bourne, 1961).

The hypothesis that the vivid imager perceives literally while the poor imager codes for retention leads to the following predictions for test: (i) An interaction between imagery type and stimulus patterning. The difference in accuracy of recall of the two stimuli which vary in patterning will be greater for the poor imagers. Regardless of the level of attribute variation, vivid imagers will be able to handle the decreasing organization of the stimuli more effectively than the poor imagers. (ii) An interaction between imagery type and level of attribute variation. The increased stimulus variability of conditions 2 and 3 makes coding more difficult for the poorer imagers. As more attributes are conjointly handled the literal perceiver will be better able to take account of the variability of the design from its basic patterning.

METHOD

Subjects

Seventy-two volunteers from psychology classes at the University of Sydney served as subjects. Independent sets of twenty-four subjects, twelve vivid and poor imagers, were randomly allocated to each of the three attribute groups. The subjects were judged to be vivid or poor imagers on the basis of their average score over all items on a shortened form of Betts' Questionnaire Upon Mental Imagery (Sheehan, 1967). The scale effectively discriminated two different types of imaging subjects. Vivid imagers reported imagery which was, at least, 'moderately clear and vivid'; poor imagers reported imagery which, at best, was 'unclear, but recognizable'.

Apparatus

Two stimuli, *A* and *B*, were rated unanimously as most and least patterned, respectively, by ten judges. The patterns were chosen from 252 black–white designs containing nine elements in all possible contrasts of 4×5 and 5×4 patterning. The basic patterning of the stimuli is given by the colour contrast of the designs in Fig. 1. Colour, shape and size of elements were arranged for conditions 1, 2 and 3, as set out in Fig. 1. Colour transparencies were prepared for screen projection of each display. The dimensions on the slides were such that, when projected on a screen at a distance of 10 ft., the diameters of large and small circles were $3\frac{1}{2}$ in. and $1\frac{3}{4}$ in. respectively.

Designs were reproducible by four sets of eighteen matt-black wooden blocks of 1 in. side. There was a set of triangles, squares, circles and diamonds. On each block a given geometric form was reproduced in red, yellow, green and blue, and the blocks in each set were equally divided into large and small shapes. The experimental room was dark except for lamp illumination of the four sets of blocks placed on a small table in front of the subject. The physical image on the screen fell ahead and slightly to the right of the subject's line of vision. A 35 mm slide projector was used with a Variac transformer, and instructions were delivered by tape-recorder.

The Kohs Block Design Test (Kohs, 1927) was used to provide a control task between perception and imaging and to familiarize the subjects with pattern construction. The imagery test which initially differentiated different types of imaging subjects contained thirty-five items pertaining to seven sensory modalities: visual, auditory, cutaneous, kinaesthetic, gustatory, olfactory and organic. All items measured a general ability to image. In the experimental setting the subjects rated their imagery on the Betts' seven-point rating scale (Betts, 1909) which ranged from 7 (no image present at all), through 5 (vague and dim) to 1 (perfectly clear and vivid).

Procedure

Prior to the experiment proper, each subject was given the imagery test to determine his imagery grouping, and received practice with the experimental task. A nine-unit geometrical pattern employing all four shapes and colours familiarized subjects with the construction and imaging of simple geometrical designs.

Stimulus *A* (or *B*) was flashed on to the screen and the subject was instructed to construct the

design by means of the blocks provided. The projected image was removed as soon as the subject matched the pattern and the experimenter disassembled the blocks while the subject worked at the Kohs block task for 40 sec. The subject was then asked to conjure up a visual mental image of the design he had just seen and to construct the pattern of his image with the blocks. The subject rated the vividness of his imagery on the rating scale. This procedure was repeated with the remaining display after an inter-trial interval of 2 min, during which the subject worked again at the Kohs block test. The order of stimulus presentation was determined randomly and inspection stimuli were projected on to the screen at a dim, supra-threshold level of illumination. All subjects were shown both stimuli.

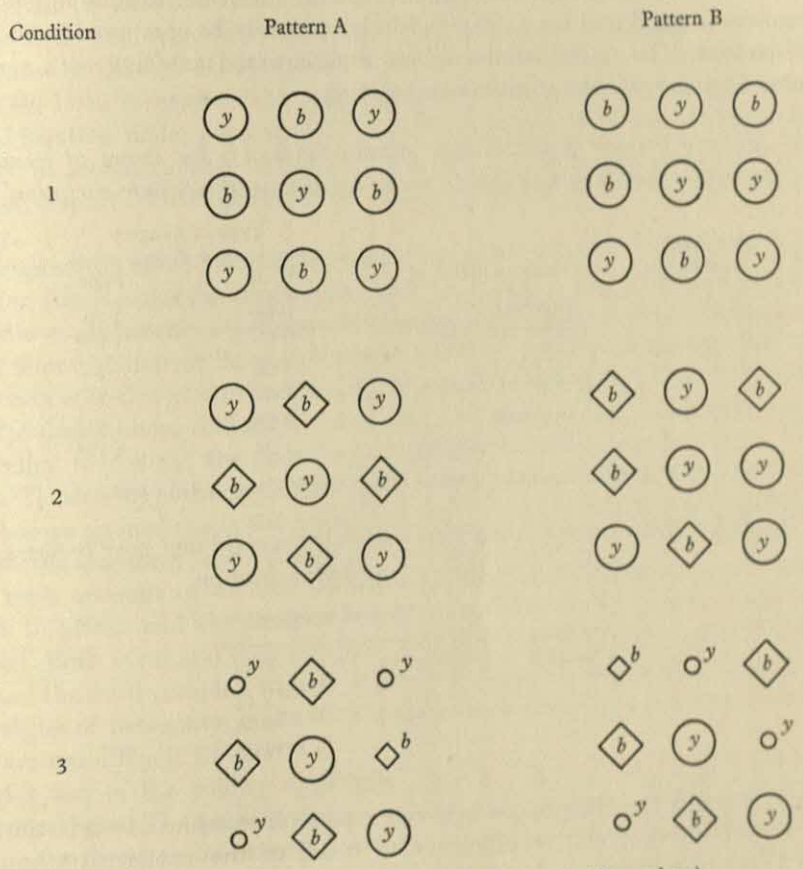


Fig. 1. Stimulus displays (b = blue colour; y = yellow colour).

Measures

Errors of location, inversion scores and vividness ratings were recorded for the subjects and analysed. Errors of location referred to the number of elements of a display in the image reproduction which were not in the same position in the stimulus pattern as perceived. Inversion scores were the number of pairwise inversions of order needed to transform the subject's ordering of elements in his imagery construction to the ordering of elements he gave in his perception construction. Small numbers of inversions indicated that an imaged pattern was quite like the perceived one, while large numbers of inversions indicated that the patterns were quite different. Details of the measure are outlined elsewhere (Sheehan, 1966).

RESULTS

Imagery groups were compared on vividness ratings given to the experimental stimuli to validate initial classification of subjects on the imagery questionnaire. A split plot analysis of variance, Lindquist's Type III model (Lindquist, 1953), indicated a significant effect for imagery type ($F = 7{\cdot}80$; D.F. 1, 66; $P < 0{\cdot}01$). Vivid imagers rated the experimental stimuli as consistently more vivid than did poor imagers over all stimuli and conditions of attribute variation. Results indicated that performances of good and poor imagers can legitimately be compared in the experimental context. The questionnaire scores differentiated subjects with respect to vividness of imagery in the experimental setting.

Table 1. *Mean difference score between stimuli* ($d(B–A)$) *for errors of location and inversions for imaging subjects under conditions of attribute variation*

Attribute condition	Measure $d(B–A)$	Type of imagery: Vivid	Type of imagery: Poor
1	Inversion	0·34	1·00
	Error of location	0·67	1·50*
2	Inversion	−1·00	1·84
	Error of location	2·42	1·41
3	Inversion	0·33	1·25
	Error of location	1·00	4·17**

* $P < 0{\cdot}05$, one-tailed t-test; ** $P < 0{\cdot}001$, one-tailed t-test.

Table 2. *Mean inversion score over stimuli for vivid and poor imagers under conditions of attribute variation*

Attribute condition	Type of imagery: Vivid	Type of imagery: Poor
1	7·25	4·42
2	4·50	8·75
3	7·04	7·05

Table 1 sets out the differences between scores on stimulus A and stimulus B. Results report on the relative efficiency of recall of the patterned stimulus for different types of imagers, and are relevant to prediction (i). Table 1 shows that difference scores were greater for subjects who imaged poorly than for those who had vivid imagery. Vivid imagers were better able to handle the complexity of the unpatterned stimulus than poor imagers.

The distributions of errors of location made parametric tests of significance impractical. There were large numbers of zero scores and wide individual differences across experimental conditions. A log transformation of inversion scores eliminated the heterogeneity of variance in four experimental groups and allowed analysis by Lindquist's Type III model of analysis of variance. Results for conditions 1 and 2 by this analysis showed a significant interaction between imagery type and variation of stimulus attributes ($F = 4{\cdot}91$; D.F. 1, 44; $P < 0{\cdot}05$). Table 2 sets out the mean inversion scores for the interaction effect which are relevant to prediction (ii). Results

indicate that good imagers behaved more poorly on condition 1 than on condition 2, while the trend was reversed for poor imagers. The poor imagers found it more difficult to handle the increased variation in stimulus attributes than did the vivid imagers. Vivid and poor imagers, though, behaved similarly on condition 3.

Discussion

Prediction (i) was that there should be a significant interaction effect between imagery type and stimulus patterning. Support for this prediction is given by examination of the $d(B–A)$ values in Table 1. Difference scores were greater for poor imagers on both measures. The only exception to this finding was the results for errors of location under condition 2. Generally, subjects with poor imagery had less difficulty in reconstructing obviously patterned displays than they had with unpatterned displays. This difference was relatively unimportant for subjects with vivid imagery.

The interaction effect between imagery type and variation of attributes confirmed prediction (ii). Results for conditions 1 and 2 indicated that vivid and poor imagers behaved very differently when more attributes were varied in the stimulus patterns. Table 2 shows that poor imagers found it far more difficult to handle the increased variation in stimulus attributes than did the vivid imagers. The increased variability made the design more difficult to code and, as stimuli were varied with respect to amenability to coding, the poor imager performed less accurately than the vivid imager. It was hypothesized that poor imagers would find reproduction more difficult since an increase in the number of attributes to be handled conjointly would decrease the accuracy and efficacy of the original coding device. Coding devices which took account of all the variation in elements would become increasingly difficult to attain and the literal perceiver would be better able to reproduce the variation. Both vivid and poor imagers performed similarly under condition 3 which presented the most complex stimuli. This result suggests that there is a point beyond which styles of perception make little difference to imagers in their recall of stimuli which are too difficult to handle.

Crucial test of the coding hypothesis requires direct manipulation of subjects' coding strategies. The present data, however, reinforce some recent evidence that encoding functions are relevant to the use and facility of visual imagery. Siipola & Hayden (1965) found a high frequency of eidetic imagery among mentally retarded children who were brain damaged. Subjects known to have considerable impairment in symbolic functioning reported exceptionally vivid imagery. Haber & Haber (1964) showed that normal subjects' actual memory for a stimulus after their eidetic imagery had faded did not differ from that of non-eidetic subjects. It was suggested that the eidetic subjects failed to take the opportunity to encode the stimulus for later recall.

In conclusion, the study indicates that organizational properties of stimuli are relevant to the retentive behaviour of visual imagers. The data are consistent with the hypothesis that differences in the retentive behaviour of imagers are related to differences in the styles of their perception. Results supported predictions made from the assertion that poor imagers code for retention while vivid imagers perceive more literally. There were some discrepancies in the results, suggesting that the hypothesis

is far from complete. Vivid and poor imagers did not differ in recall under condition 3 and not all measures yielded consistent results. The data indicate, however, that encoding processes may play a complex role in determining the accuracy of retention for imaging subjects. Further work is needed to investigate the parameters which determine the extent to which the vivid imager behaves differently from the poor imager under conditions of increased stimulus complexity.

This study was conducted as part of a doctoral programme of research on visual imagery. It was supported by U.S. Public Health Service N.I.M.H. Grant M-3950. The author is indebted to Dr J. P. Sutcliffe, Principal Investigator for the above grant, for the stimulation of his ideas and his help throughout the study.

REFERENCES

AVERBACH, E. & SPERLING, G. (1961). Short-term storage of information in vision. In *Fourth London Symposium on Information Theory*, ed. C. Cherry. London: Butterworths.

BATTIG, W. F. & BOURNE, L. E. (1961). Concept identification as a function of intra- and inter-dimensional variability. *J. exp. Psychol.* **61**, 329–33.

BETTS, G. H. (1909). The distribution and functions of mental imagery. *Columb. Univ. Contrib. Educ.* **26**, 1–99.

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

GARDNER, R. W. (1953). Cognitive styles in categorizing behaviour. *J. Pers.* **22**, 214–33.

HABER, R. N. (1964). Effects of coding strategy on perceptual memory. *J. exp. Psychol.* **68**, 357–62.

HABER, R. N. & HABER, R. B. (1964). Eidetic imagery: 1. Frequency. *Percept. mot. Skills* **19**, 131–8.

KOHS, S. C. (1927). *Intelligence Measurement*. New York: Macmillan.

LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton-Mifflin.

MACKWORTH, J. (1963). The duration of the visual image. *Can. J. Psychol.* **17**, 62–81.

SHEEHAN, P. W. (1966). Functional similarity of imaging to perceiving: Individual differences in vividness of imagery. *Percept. mot. Skills* **23**, 1011–33. (*Monogr. Suppl.* 6–V23.)

SHEEHAN, P. W. (1967). A shortened form of Betts' Questionnaire Upon Mental Imagery. *J. clin. Psychol.* **28**, 386–9.

SIIPOLA, E. M. & HAYDEN, S. D. (1965). Exploring eidetic imagery among the retarded. *Percept. mot. Skills* **21**, 275–86.

(*Manuscript received* 6 *December* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 253–259
Printed in Great Britain

AN ANALOGUE OF EXTRAVERSION AS A DETERMINANT OF INDIVIDUAL DIFFERENCES IN BEHAVIOUR IN THE RAT

By E. WELDON
Institute of Psychiatry, University of London

The investigation was concerned with an analogue of extraversion–introversion in rats as precursor to a selective breeding experiment. Forty Maudsley reactive and forty Maudsley non-reactive rats were tested in both the Mowrer box and in a spatial alternation T-maze. It is argued that each of these test situations gives a measure of extraversion–introversion through the involvement of reactive inhibition in each case (cf. Eysenck, 1955, 1957). Both measures exhibited strain differences; but comparison of the scores of individual subjects also suggested individual differences that can be attributed to differences in stimulus hunger.

Eysenck (1955, 1957) formulated a theory of personality in terms of orthogonal dimensions, of which the two principal dimensions are neuroticism and extraversion–introversion. He linked extraversion–introversion to Hull's concept of reactive inhibition (Hull, 1943), which is postulated to account for individual differences in behaviour in various experimental situations; for example, in the tendency to alternate (House, 1956), in the learning of integrative responses (Eysenck, 1963*a*), and as a measure of the excitatory state in rats (cf. Eysenck, 1963*b*; Sinha, Franks & Broadhurst, 1958). The investigation reported here consisted of a comparison of the measurements obtained from two experimental test situations, each of which involved reactive inhibition and hence an analogue extraversion.

In the first of these situations a Mowrer box was used, in which the individual differences in the learning of an integrative response are taken to be due to individual differences in conditionability. It has been shown with human subjects that introverts condition better than extraverts (Franks, 1957), so that the differences in conditionability in the Mowrer box may be considered to lie on a scale similar to the extraversion–introversion scale. The other test situation involved the Sinha alternation maze, which was taken to measure reactive inhibition and to be independent of emotionality. It was hypothesized that those subjects that generated reactive inhibition quickly would alternate more than subjects that generated reactive inhibition more slowly.

In comparing the individual scores in the two test situations a correspondence between them would be strong evidence for a dimension in rats analogous to that of extraversion–introversion in humans. Such a dimension is postulated to be orthogonal to neuroticism, for which an analogue in rats has been described by Broadhurst (1960) as 'emotionality'; he established the Maudsley reactive (MR) and Maudsley non-reactive (MNR) strains by means of selective breeding for emotional elimination in the modified Hall open-field test.

The use of Maudsley reactive and Maudsley non-reactive rats in this investigation provided two groups of subjects known to differ in emotionality. If these two groups also differed from each other significantly in any of the test situations, such strain differences might indicate that the particular tests were not independent of emotionality and therefore that the analogue dimension of extraversion was not orthogonal to the analogue dimension of neuroticism.

METHOD

Subjects. The subjects used were 40 rats each of the Maudsley reactive (MR) and Maudsley non-reactive (MNR) strains (20 males and 20 females in each case) of the 22nd generation of selection, all between 113 and 176 days old on the first day of the experiment. All subjects were reared under the standard conditions described by Broadhurst (1960). In their home cages they had access to water at all times but feeding was for a period of 2 hr daily, commencing 1 hr after each experimental session. The feeding schedule was established two weeks prior to the start of the experiment.

Apparatus. The Mowrer box used was built from the description given by Mowrer & Miller (1942). It is essentially an alley with an electrifiable floor, having a food-pellet dispenser and trough at one end, and an associated buzzer. The Sinha maze was an enclosed spatial alternation T-maze, as described by Sinha *et al.* (1958). It differs from other mazes in its arrangement. A subject running the maze has to make 3, 2, 1 or 0 forced left- or right-hand turns before reaching the choice-point; depending on the starting position (3, 2, 1 or 0 respectively) selected.

Procedure. The procedure used in the Mowrer box was based on that of Mowrer & Ullman (1945) and Eysenck (1963*a*). Rats on a 22 hr fasting schedule were trained to eat a pellet of food presented at the termination of a buzzer sound of 2 sec duration, on ten trials spaced at 1 min intervals on ten consecutive days. Following this, 100 similar trials were given when the duration of the buzzer was extended by 3 sec beyond the presentation of the food pellet. If the subject had not taken the food pellet by the 55th sec of the trial, the food was tipped out of the trough away from the reach of the subject.

A 'rule' was then introduced that it was 'incorrect' to eat whilst the buzzer was sounding, i.e. the previous regime was changed. When the food was picked up before the termination of the buzzer, the duration of the buzzer was extended by a further 9 sec (14 sec total duration) and an electric shock (0·3 mA) applied to the subject's feet for the last 2 sec of this extended period. Under this routine, the subjects could respond in three ways. They could eat before the buzzer stopped, eat after the buzzer stopped, or not eat at all; these specific responses were tentatively identified as corresponding respectively to psychopathic (delinquent) behaviour, integrative behaviour, and dysthymic (neurotic) behaviour.

The procedure used with the Sinha maze was adapted from that described by Sinha *et al.* (1958); since in the present study the interest centred on individual differences in behaviour, each subject was started from every starting position. Each subject therefore was given 1 day of preliminary adaptation training, followed by training from starting position 0 on the second day, test runs from starting position 0 on the third day, training from starting position 1 on the fourth day, test runs from starting position 1 on the fifth day, and so on for starting position 2 and 3; the whole procedure taking 9 days in all.

For the preliminary training, all the temporary blocks in the alleys were removed and the troughs in the two goal arms of the T filled to the brim with pellets. The subjects were placed into the maze at starting position 0 and left in the apparatus for 10 min. On subsequent days during both training and test runs, the subjects were given 'spaced' trials individually, in order to minimize the building of reactive inhibition. On each day when subjects were trained to run from particular starting positions, temporary blocks were positioned in the alley so that the starting position and the particular goal arm of the T constituted the extreme ends of a series of consecutive straight alleys without any choice points.

In order to avoid establishing position preferences in the subjects (i.e. right or left), both the side of the starting position and the side of the goal arm used in any particular trial were randomized respectively by the use of two separate Gellerman series (Gellerman, 1933). Such series ensured that not more than two consecutive runs were in any one direction. The double randomization procedure (for starting positions and goal arms) cut out the influence of orientating cues leading to position preference, and built up equal habit strengths for running to both goal arms from each of the various starting positions. One food pellet was placed in the appropriate trough before each run.

Two measures were taken: the number of alternations and the latency (i.e. the time between the introduction of the subject into the apparatus and arrival at the food trough). Each subject received ten trials.

On every day when subjects were given 'test' runs, each subject received six 'test' or 'free choice' runs, each of which was followed by a 'forced choice' run. There were in all twelve runs for each subject. The temporary blocks were positioned as during the previous day's training runs, except that on 'free choice' runs a choice point was created in the T by the removal of the temporary block from the second goal arm. The appropriate left or right starting position was selected as before by a modified Gellerman series. But the goal side blocked on the forced choice runs was in each case determined by the subject's own choice on the previous 'free choice' run, such that each subject was forced away from the side previously chosen. This ensured that prior to each test run, the subject had eaten an equal number of times in each of the two goal arms; it was therefore an additional precaution against the development of position habits.

Results

Two kinds of statistical analysis were required. First, each experimental situation had to be examined for effects due to strain and sex differences in the subjects, in order to eliminate or control these factors in the subsequent analyses. Following this, the two experimental situations were compared with each other.

Table 1. *Analysis of variance of scores in Mowrer box during training period*

		Frequency score (e')		Latency score (l')	
Source	D.F.	M.S.	F	M.S.	F
Strain	1	0·034	3·04	0·450	36·08**
Sex	1	0·007	0·75	0·012	0·96
Strain × sex	1	0·001	0·09	0·033	2·64
Residual	76	0·011		0·012	
Total	79				

** $P < 0·01$.

During the training period for the Mowrer box, two measures were taken, following Eysenck (1963*a*). The first was the proportion of the number of trials when the subject ate the food pellet during the 100 trials (10 trials on each of 10 days), called the 'e' score. The second was the cumulative latency score over the ten trials of the last 'training' day called the 'l' score. Because the distributions of both these measures were skewed, the scores were transformed to $e' = \sin^{-1}\sqrt{e}$ and $l' = \log l$. Analyses of variance were applied to the transformed scores. The results (see Table 1) show a significant strain effect ($P < 0·01$) in respect of the latency score but the analysis of the e' scores shows no significant results.

During the test period for the Mowrer box three kinds of scores were obtained, i.e. a scores being the frequency of the 'integrative' response, b scores being the frequency of the 'delinquent' response, and c scores being the frequency of the 'neurotic' response. Each of these scores was examined by separate analyses of variance (see Tables 2–4), strain and sex being taken as main effects. The analyses show that the main effect of strain on c scores and the strain × sex interaction for both a and b scores were significant at the 5 % level. The MNR males and MR females had about the same number of a and b scores, but the MNR females made more b responses than the MR males (see Table 5).

The T-maze situation provided an alternation frequency score and a latency score. But for purposes of statistical computation, derived scores were obtained by summing the number of alternations in respect of the six test runs for each starting position.

Similar derived scores were obtained for the corresponding latencies, L. Because the distribution of L was skewed, logarithmically transformed scores ($L' = \log L$) were used in the computations, in order not to violate assumptions of normality. An analysis of variance was made for each of the two scores. The results (see Table 6) showed a significant strain effect ($P < 0{\cdot}01$) for both alternation and latency scores. The MR subjects tended to alternate less and have larger latencies than the MNR subjects.

Table 2. *Analysis of variance of Mowrer box 'integrative' responses scores, a*

Source	D.F.	Mean square	F	P
Strain	1	0·125	0·01	N.S.
Sex	1	0·445	0·08	N.S.
Strain × sex	1	78·125	4·52	< 0·05
Residual	76	17·320		
Total	79			

Table 3. *Analysis of variance of Mowrer box 'delinquent' response scores, b*

Source	D.F.	Mean square	F	P
Strain	1	54·080	3·16	N.S.
Sex	1	4·805	0·28	N.S.
Strain × sex	1	76·880	4·49	< 0·05
Residual	76	17·132		
Total	79			

Table 4. *Analysis of variance of Mowrer box 'neurotic' response scores, c*

Source	D.F.	Mean square	F	P
Strain	1	50·501	5·13	< 0·05
Sex	1	1·532	0·16	N.S.
Strain × sex	1	0·362	0·04	N.S.
Residual	76	9·851		
Total	79			

Table 5. *Mean frequencies of the a, b and c responses in the Mowrer box by male and female Maudsley reactive (MR) and non-reactive (MNR) rats*

Strain	Sex	*a*	*b*	*c*
MNR	♂	64·0	24·7	11·3
MNR	♀	56·9	32·4	10·7
MR	♂	58·3	25·7	16·0
MR	♀	63·5	21·0	15·7

The T-maze and Mowrer box situations were compared with each other by means of a number of correlations. Both the alternation score and the log transformed latency score of the T-maze were in turn correlated with the Mowrer box *a*, *b* and *c* scores. The respective correlation coefficients are shown in Table 7; none exceeded 0·21.

All the subjects, for each strain separately, were ranked according to their individual T-maze alternation scores, and the upper 25 % of the subjects compared with the lower 25 % in respect of the mean values of their *a*, *b* and *c* scores in the Mowrer box. Differences between the subjects in these two quartiles were then tested for significance

using the t-test (see Table 8). Significant differences at the 5% level were found for the c scores in the MNR strain and the b score in the MR strain, but there were no other significant differences between the upper and lower quartiles of each strain.

Table 6. *Analyses of variance of* T-*maze scores*

		Alternation score		Latency score (L')	
Source	D.F.	M.S.	F	M.S.	F
Strain	1	59·513	29·91**	11·463	167·88**
Sex	1	4·050	2·04	0·002	0·02
Strain × sex	1	0·200	0·10	0·046	0·68
Residual	76	1·990		0·068	
Total	79				

** $P < 0·01$.

Table 7. *Correlation coefficients between Mowrer box scores and* T-*maze scores*

	T-maze	
Mowrer box	Alternation score	Log latency score
a score	0·07	0·03
b score	0·19	0·14
c score	−0·21	0·16

Table 8. *Comparisons between the means of two groups of subjects in respect of their a, b and c Mowrer box scores; the two groups consisted of the subjects that were in the respective upper and lower quartiles of the ranked* T-*maze alternation scores*

		Mowrer box means			t between upper and lower quartiles		
Strain	T-maze alternation range scores	a	b	c	a	b	c
MNR	Lower quartile 11–16	58·4	21·1	20·4	0·84	1·17	1·91*
	Upper quartile 21–24	66·5	29·8	3·7			
MR	Lower quartile 10–13	67·1	13·2	19·6	0·74	2·13*	0·95
	Upper quartile 16–22	61·9	24·2	14·1			

* $P < 0·05$.

Discussion

The specific aim of this experimental study was to test the prediction that there would be a correspondence in the individual differences in behaviour of rats in two test situations. However, it was necessary to examine the results of each of the test situations for contamination due to sex and strain factors, since an analogue scale for extraversion should by definition be independent of such factors.

Examination of the results (Table 6) showed that for the T-maze situation there was a highly significant effect of strain on both alternation and latency scores ($P < 0·01$), but no effect of sex. The MR strain had smaller alternation and greater latency scores than the MNR strain. This was contrary to the findings of Sinha *et al.* (1958), who had found a sex difference but no strain differences. There were differences in the experimental procedures between the two studies, however, which

necessitated differences in the respective training procedures. For example, in the present study, training and test days were interspersed because every subject started in turn from each of the starting positions, whereas each of Sinha's subjects started always from the same starting position. Again, the respective amounts of training were not quite comparable. It could be argued, for example, that in Sinha's experiment the subjects had attained greater familiarity with the apparatus before the beginning of the testing period, and that there had in general been a better adaptation to the experimental situation.

The results obtained during the training period for the Mowrer box situation (Table 1), in common with those of Eysenck (1963*a*), showed no differences attributable to sex and a significant strain effect in respect of the latency score (l'). Eysenck (1963*a*) found a significant strain effect for the frequency score (e') at the 5 % level. In the present study, the level of significance of the strain effect fell only slightly below 5 %, so that the discrepancy was small and probably attributable to a sampling error.

The analyses of variance of the frequency scores for the testing period (Tables 2–4) showed a significant strain effect in the *c* (not eating) score only. This result would seem to be in accord with Eysenck (1963*a*), who reported a significant mode × strain interaction. Further examination of the results, comparing the T-maze alternation score with the Mowrer box *a*, *b* and *c* scores (see Table 8), was carried out for the two strains (MR and MNR) separately, because of the significant strain effects found in the analyses of variance for the T-maze and Mowrer box situations. The results showed that the MR strain subjects tended to have larger *c* scores and smaller *b* scores than the MNR strain subjects. There were also significant differences between the upper and lower quartiles (ranked on the T-maze alternation scores) in respect of the *b* (eating before buzzer stops) score of the MR strain and in respect of the *c* (not eating) score of the MNR strain, but none in respect of the *a* (eating after buzzer stops) scores of either strain. This suggests that *a* and *c* types of behaviour are typical of the MR strain, and *a* and *b* types of behaviour of the MNR strain. If *c* responses are interpreted as 'latencies greater than 55 sec', then high *c* scores, which are found to be typical of the MR strain in the Mowrer box, represent high latency scores. This finding is compatible with, and confirms, that of a strain difference in the T-maze latency scores.

All the correspondence between the T-maze and the Mowrer box situations can be explained on the basis of emotionality, because both situations were shown to be sensitive to strain. However, whilst an explanation in terms of the strain difference in the Mowrer box may account for the preponderance of *b* and *c* responses by subjects of the respective strains (i.e. *b* responses by MNR, *c* responses by MR), the differences between the scores within the respective upper and lower quartiles can be attributed to a factor other than emotionality (a factor that need not necessarily be orthogonal to emotionality). This interpretation lends support to the explanation of behaviour in terms of growth and decay of reactive inhibition, which in turn is consistent with an analogue dimension of extraversion–introversion. If the correlations between the scores on the two tests had been high, then both tests could be regarded as measures of reactive inhibition. The correlations were in fact rather low (see Table 7). It seems possible that if the two tests measure reactive inhibition the highly significant strain

effects would mask this. In order to avoid such contamination of the common factor which has been postulated, larger numbers of subjects from strains heterogeneous for emotionality would have to be used.

It has been shown that those rats which obtained higher alternation scores also tended to have high Mowrer box *b* scores (eating before buzzer stops). Now these *b* scores could be regarded as anticipatory responses. Such anticipation may be due to an exploratory drive; a high level of exploratory drive in a subject would result in a large number of anticipatory responses. Exploration in the Mowrer box involves ambulation and the animals sniffing the apparatus, food trough, food, etc. This behaviour is likely to increase the chance of food being discovered and eaten, and therefore to result in an apparent increase in observable anticipatory responses.

It could be suggested that rats showing a high degree of anticipatory behaviour, and also a high frequency of alternation responses, exhibit a strong stimulus hunger. This term has been suggested by Eysenck (1963*b*) to describe the tendency of extraverted humans to seek strong and continual sensory stimulation, whereas introverts tend to shun such stimulation. Hence, rats having a strong stimulus hunger could be said to show behaviour analogous to that of human extraverts.

Such an interpretation of behaviour in rats in terms of stimulus hunger and reactive inhibition lends support to the view that a dimension analogous to extraversion–introversion exists in rats. However, it will be necessary to carry out further investigations, using strains of rats heterogeneous in respect of emotionality, in order to confirm the existence of this dimension, before it can be used as a selection criterion for breeding.

The author is indebted to Professor H. J. Eysenck for his suggestions and encouragement with this investigation. Thanks are also due to Professor P. L. Broadhurst, Dr H. C. Holland and Dr A. E. Maxwell for their advice and helpful criticisms. The investigation was supported by a grant from the Bethlem Royal Hospital and Maudsley Hospital Research Fund.

References

Broadhurst, P. L. (1960). Experiments in psychogenetics: applications of biometrical genetics to the inheritance of behaviour. In *Experiments in Personality*, ed. H. J. Eysenck, vol. 1, pp. 3–102. London: Routledge & Kegan Paul.

Eysenck, H. J. (1955). A dynamic theory of anxiety and hysteria. *J. ment. Sci.* **101**, 28–51.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge & Kegan Paul.

Eysenck, H. J. (1963*a*). Emotion as a determinant of integrative learning: an experimental study. *Behav. Res. Ther.* **1**, 197–211.

Eysenck, H. J. (1963*b*). *Experiments with Drugs*. London: Pergamon Press.

Franks, C. M. (1957). Personality factors and a rating of conditioning. *Br. J. Psychol.* **48**, 119–26.

Gellerman, L. W. (1933). Chance order alternating stimuli in visual discrimination experiments. *J. genet. Psychol.* **42**, 206–8.

House, B. J. (1956). The relationship of place and turn alternation to reactive inhibition. *J. genet. Psychol.* **89**, 183–97.

Hull, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century-Crofts.

Mowrer, O. R. & Miller, N. E. (1942). A multi-purpose learning-demonstration apparatus. *J. exp. Psychol.* **31**, 163–70.

Mowrer, O. R. & Ullman, A. D. (1945). Time as a determinant in integrative learning. *Psychol. Rev.* **52**, 61–90.

Sinha, S. N., Franks, C. M. & Broadhurst, P. L. (1958). The effect of a stimulant and a depressant drug on a measure of reactive inhibition. *J. exp. Psychol.* **56**, 349–54.

(*Manuscript received* 21 *December* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 261–270
Printed in Great Britain

SOME PROBLEMS IN REPERTORY GRID MEASUREMENT

I. THE USE OF BIPOLAR CONSTRUCTS

By J. M. M. MAIR
Academic Department of Psychiatry, Middlesex Hospital Medical School, London, W.1

Kelly (1955) assumed that constructs were bipolar, and developed the repertory grid as a method for measuring links between construct dimensions. In most grid applications implicit assumptions seem to be made that elements not chosen in grid sorts using one pole of a construct dimension will be chosen in sorts based on the other pole, and that relations between constructs derived from grid sorts using one pole of a construct will be the converse of the relations derived from sorts based on the other pole. The legitimacy of these assumptions was examined experimentally with the co-operation of ten male medical students acting as subjects. Considerable *overlap* in element choices based on opposite poles of five construct dimensions appeared for each subject, and the practical effects of this overlap on various types of grid scores were demonstrated. Suggestions for modifications in grid usage, and for further exploration of construct functions, are discussed.

Kelly (1955) suggested that people make sense of themselves and of other events in the world around them by erecting some kind of personally organized hierarchy of interpretations of these events. He regarded each interpretation or construct as an abstraction which subsumes some aspect of a repeated event or set of events of which the person is aware. Each construct applies only to a limited number of events in any person's life, and each is bipolar. Kelly assumed that the relation between the two poles of any construct is essentially one of contrast. It follows from this idea that one person cannot hope to understand fully the meaning of any construct used by another unless the implied contrast involved for that other individual is appreciated.

The repertory grid method devised by Kelly for measuring various aspects of the relations or linkages between constructs depends heavily on many of the original theoretical assumptions about the nature of constructs. He thus laid stress on the importance of eliciting the personal construct dimensions used by the person being tested; on having that person specify both poles of each construct dimension; and on ensuring that all the elements or events in the grid could be subsumed within the range of convenience of every construct.

The primary concern here is not to examine the usefulness of assuming bipolarity in construct usage, but to investigate some problems in translating the notion of bipolarity into terms of practical grid measurement. At present many grids are administered for various purposes in which constructs are supplied by the experimenter or clinician rather than elicited from the subject or patient. This in itself is not necessarily unsatisfactory since, no matter what the verbal label given, the construction placed upon it is personal. However, this personal use of labels seems largely to be ignored in testing since the subject is seldom asked to specify the contrast pole he is using. Even in situations where constructs are elicited from the subject, and both poles are specified, the experimenter seldom asks for elements

to be sorted on the basis of both poles. An assumption seems to be made that the element selection resulting from the use of one pole would be the converse of that for the other. Although this assumption might not be accepted when made explicit by those familiar with grid measurement, those less familiar with such methods may still be liable to make inferences about the meaning of some of their results which the present measurement technique does not justify. The point is not a mere methodological detail, since misunderstanding of grid limitations could readily lead to mistaken interpretations of results, and could affect the future course of some research or treatment programme.

The aims of the investigation were: (*a*) to illustrate the personal nature of pole contrasts in terms of the words used to index contrasts; (*b*) to examine the legitimacy of the assumption that the elements not chosen on the emergent pole of a construct are those which will be subsumed by the implicit pole; (*c*) to examine the importance of the above assumption for the measurement and interpretation of construct relation scores. Possible explanations of the results are considered, and their implications discussed.

METHOD

Subjects. Ten male medical student volunteers.

The grid. A split-half form of the grid method (Bannister, 1960; Mair, 1966) with twenty elements was used. The elements were the names of people known well by the subjects, each subject being asked to include people he liked and ones he disliked, males and females. Four adjectives were supplied to each subject, and each was asked to specify what for him was the opposite of each characteristic, as it would apply to the description of people. The adjectives supplied were: KIND, INTELLIGENT, GENEROUS and ENERGETIC. The terms MOST LIKE MYSELF, LEAST LIKE MYSELF, LIKE I WOULD LIKE TO BE, LIKE FATHER, LIKE MOTHER and KNOW BEST were also included.

Administration. The grid was administered in the usual manner except that grid sorts were made in terms of each of the four adjectives *and* of each of their contrasts, in addition to sorts for the other six terms. On each sort the subject was asked to select the ten people (elements) who in his opinion showed the characteristic involved most clearly. The experimenter entered ticks in the grid blank to represent the subject's element choices on each grid sort. When a subject made his first sort for a contrast-pole adjective he was reminded that he should focus his attention on the particular term under consideration, and not base his decisions on a recollection of earlier choices made for the alternative pole. Each was told that he might sometimes find it appropriate to choose all ten elements at one pole which were excluded on the contrast-pole sort, and sometimes express his own views most adequately by selecting or excluding certain elements in both pole sorts.

Scoring. The scoring procedure has been described fully by Bannister (1960). In a split-half grid form with twenty elements there are eleven possible matching scores between constructs, ranging by steps of two from 0 to 20. The statistical significance of matching scores of different sizes can be calculated by the usual method of binomial expansion. With twenty elements, $P = 0{\cdot}12$ for matchings of 6 and 14, and $0{\cdot}01$ for matchings of 4 and 16. In this study, matchings of 14 or higher have been accepted as indicators of *positive* associations between constructs; matchings of 6 and less as indicators of *negative* associations; and those between 8 and 12 as representing essentially *chance* relations.

RESULTS

Personal contrasts in construct usage

The variety of words used to specify the personal contrasts to each of the four supplied adjectives is shown below.

KIND—unkind, thoughtless, inconsiderate, malicious, mean, unpleasant.

INTELLIGENT—stupid, thick, dull.
GENEROUS—mean, selfish, stingy, close.
ENERGETIC—lazy, lethargic, listless, sleepy.

Overlap between opposites

Matching scores were calculated for each person between the four supplied adjectives and their elicited contrasts, and between LIKE SELF and UNLIKE SELF. These matchings (five for each person) are shown in Table 1. If the assumption that elements not chosen at one pole of a construct will be chosen at the other is correct, then very low matchings (0 or 2) should appear throughout. Higher scores may cast doubt on the adequacy of the assumption. Only 22% of the matchings were 0 or 2, and 54% were as high as 6 or 8. None was above 8. Each subject had at least one matching which was 6 or above.

Table 1. *Matching scores between five polar opposites*

Subjects	KIND	INTELLIGENT	GENEROUS	ENERGETIC	LIKE SELF
1	4	0	2	6	4
2	6	8	2	8	2
3	4	0	6	4	8
4	4	6	4	6	8
5	2	6	6	4	8
6	6	6	4	6	4
7	6	0	6	2	0
8	4	2	6	6	8
9	6	4	8	6	6
10	4	2	6	8	6

Importance of overlap for the interpretation of grid scores

(*a*) *Converse relations*

For each person and each of the five pole contrasts, selections of elements made at each pole were matched with those for the twelve other constructs in the grid (excluding its contrast). For each individual there were therefore 120 matchings, providing sixty matching score comparisons between pole contrasts.

If the assumption being tested were valid, or if the overlap scores noted in Table 1 were unimportant, then all matchings between pole contrasts and the given set of constructs should be converse or complementary. In this instance converse relations were assumed to exist under the following conditions—when the emergent pole matching was positive and the implicit pole matching was negative; when the emergent pole matching was negative and the other positive; and when both were at a chance level. This provided a fairly gentle test of the assumption since differences in degrees of relation between constructs were ignored, provided a generally converse relation between pole contrasts existed. Thus the emergent pole of a construct might provide a matching of 18 (i.e. a deviation of 8 above the absolute chance score of 10) with another construct, while its contrast (the implicit pole) gave a matching of 6 (i.e. a deviation of 4 below the absolute chance score) with the same construct, and these would be accepted as indicating a converse relationship.

Even in this situation 32% of all matchings of pole contrasts with other constructs

were not converse. The range of individual variation in non-converse relations was 17–55%. If a more precise definition of converse relations were to be used, involving identical deviations from matchings of 10, more than twice this percentage of non-converse relations would appear. Thus in a third of the 600 comparisons involved, the assumption that construct relation scores derived from matching element selections made at the implicit pole of a construct would be the converse of those at the emergent pole, would be in error.

(*b*) *Intensity of construct relations*

Although the assumption of converse relations between opposite poles of a construct and other constructs may sometimes be inappropriate when overlap scores are present, it is still possible that more formal structural or mathematical aspects of construct relations might not be affected. Intensity, one of the most commonly used structural measures, has been examined.

Intensity scores indicate something of the overall degree of association between a given set of constructs. Intensity scores were calculated by summing the deviations from ten of the matchings of selections at each pole of each of the five sets of opposites with the twelve other constructs in the grid (regardless of sign) for each individual. Comparisons of the distributions of intensity scores at each pole were made by using the Wilcoxon test for matched pairs (Siegel, 1956).

Table 2. *Individual differences in intensity of matching of pole opposites*

Subject	Differences	Wilcoxon comparisons		
		n	T	P
1	LIKE SELF > UNLIKE SELF	9	4·5	< 0·05
2	STUPID > INTELLIGENT	11	3	< 0·01
	ENERGETIC > LAZY	10	5	0·02
3	GENEROUS > SELFISH	9	6	0·05
6	GENEROUS > MEAN	9	2·5	< 0·02
9	THICK > INTELLIGENT	8	0	0·01
10	GENEROUS > MEAN	10	7	< 0·05
	UNLIKE SELF > LIKE SELF	9	6	0·05

The terms GENEROUS and INTELLIGENT provided statistically significant differences in intensity scores for matchings derived from each pole. For the group, GENEROUS showed higher intensity of matchings with other constructs than did its opposite ($n = 9$; $T = 3$; $P = 0{\cdot}02$). INTELLIGENT showed a generally less intense degree of matching with other constructs than did its opposite ($n = 7$; $T = 0$; $P = 0{\cdot}02$).

Calculations were also made to check on possible differences in overall degree of relation at each pole of each of the five sets of opposites within individual subjects. Again Wilcoxon's test (two-tailed) was used, and five comparisons made for each subject. Six subjects provided significant differences in intensity of matchings of pole opposites on at least one construct. These differences are shown in Table 2.

The narrow range of scores available in these within-individual comparisons made difficult the attainment of significant differences. Still, the constructs showing the group differences again appeared most frequently (i.e. GENEROUS and INTELLIGENT).

The results for LIKE SELF/UNLIKE SELF comparisons for subjects 1 and 10 were particularly interesting in that for the former LIKE SELF showed greater intensity of matchings with the other constructs than UNLIKE SELF, while for the latter more marked relations appeared when the selection of elements for UNLIKE SELF was used.

In view of the fact that Bannister (1960, 1962) has repeatedly used a more generalized form of intensity score than that considered here, it is perhaps important to note that when a similar assessment was made (i.e. by summing across constructs as well as subjects) no difference in intensity of matchings of constructs with the emergent and implicit poles appeared. This would not seem surprising in the present instance, since, for the two constructs which did provide significant group differences, the emergent pole showed more intense relations in one case (GENEROUS), and the implicit pole more intensity in the other (INTELLIGENT).

Possible explanations of overlap scores

There are many reasons why overlap in element selections based on opposite poles of constructs might appear. Some of the possible explanations were mentioned by students commenting on the task during the testing session. One possibility would seem to be that the elements involved were very complex and open to many interpretations. People obviously act differently at different times and in different situations. If a construct is used more to subsume actions, intentions, potentials and the like, rather than to summarize the total person, it is quite possible for any person to be seen as demonstrating opposing characteristics. Even if used to describe the whole person, both ends of many constructs could be applied to one person at different times or in different situations.

It is also possible that the split-half method used here might have contributed to overlaps. If some of the elements were outside the range of convenience of a construct, then, since the subject was required to select ten elements, he might include one or two people whom he considered generally STUPID to complete his selection for INTELLIGENT. So also when he made his selection for STUPID, since he could not meaningfully describe some of the elements, or did not know enough about them to make a decision, he might prefer to use elements already chosen at the opposite pole. Another possibility is that some of the dimensions involved might only provide crude discriminations for some subjects. Thus some might find that they could make no meaningful distinctions between certain elements on some dimensions. In this situation, when selections were made at each pole, choices between some elements which fell in the middle range might be virtually a chance procedure, since no real distinctions could be made between them.

Further explanations derived from direct questioning of the students, in the course of a subsequent study, about the bases for some of their element choices. Here attention was focused only on sorts involving LIKE SELF and UNLIKE SELF. Subjects were shown the names of people they had included in both the LIKE SELF and UNLIKE SELF sorts, and those they had excluded from both, and asked to explain their decisions. In general it seemed that people were excluded from both sorts because the subject had had no real contact with the persons and so did not know them well enough; or the people had made some big impact on the subject (e.g. mother, best

friend, girl friend), raising conflicting ideas or feelings which he could not yet resolve; or the subject was in the course of changing his ideas about some person, but had not yet reached a stable present assessment.

Elements were included as both LIKE and UNLIKE SELF because they were seen as similar in some respects but different in others. Perhaps the most common explanation here was that people were seen as similar to self in having comparable home or educational backgrounds, interests or potentialities, but different in the ways they responded to, used or developed these.

DISCUSSION

Kelly accepts as part of the assumptive structure of his theory that constructs are bipolar and are used basically in a dichotomous fashion. These conceptions, therefore, do not need to be tested within the framework of his theory, and the present discussion is not concerned with questioning the possible usefulness of the assumptions. Its concern is rather with exploring some problems arising from the assumption of bipolarity and some implications of these for measurement.

Kelly has emphasized the personal nature of construction, even when there is apparently great communality in the language used to express such constructions. The contrasts between poles of constructs used by individuals often differ considerably from those which would seem to be dictated by the logic of the public language. Resnick & Landfield (1961) have given substance to this point in their distinction between 'logical' and 'peculiar' constructs. In the present study, although none of the contrast poles elicited from subjects were very peculiar, the diversity of labels used in this small homogeneous group suggests that the elements that would be subsumed by the various dimensions would not always be identical. Both studies seem to demonstrate enough variety in pole contrasts to encourage more attention to this problem in psychological measurement generally, and grid measures in particular.

Only one previous study has examined directly the overlaps in element choices resulting from the use of pole opposites (Bridges, 1965); considerable overlaps appeared for two contrasts (GOOD/BAD and SINCERE/INSINCERE) for small groups of normal subjects, for thought-disordered schizophrenic patients and for non-thought-disordered schizophrenic patients. For the SINCERE/INSINCERE comparison, the normal subjects' overlap scores did not differ significantly from thought-disordered scores. One limitation of this study, mentioned by Bridges, was that both poles were supplied to the subjects, and some of the overlaps might have been due to difficulties in adjusting personal constructions to the prescribed dimensions. The overlaps noted there were, however, greater than could reasonably be attributed to this factor alone.

In the present study, though the initial adjectives were supplied, the contrast poles used by each subject were elicited so that subjects could be allowed an opportunity of using more personally meaningful dimensions. In addition to this, more familiar elements were used (people known by the subjects rather than photographs of unknown faces), and a greater number of dimensions was involved. Overlaps still appeared, and matching scores between pole opposites, though generally low,

were considerably greater than zero for most subjects. It might be objected that the instructions given to subjects when they made their first sorts using contrast-pole adjectives could have encouraged carelessness and loose thinking. There was, however, no evidence of this in the test situation or in subsequent reports by those participating. Preliminary trials of the testing procedure had rather indicated the need for such instructions to counter a strong tendency for subjects to interpret the task as a test of logical consistency.

The finding that pole opposites can be related at a chance level is disturbing on a number of counts.

(1) The present custom in grid measurement of using only one pole for grid sorts may result in many mistaken inferences about construct relations. At present, for example, if selections based on one construct pole, like ENERGETIC, are very similar to those for INTELLIGENT, it is very tempting to assume that LAZY relates positively to STUPID, and negatively to both INTELLIGENT and ENERGETIC. The results in 3(*a*) indicate that this type of assumption may frequently be unjustified. Such an assumption may often be made in the clinical use of grids to understand something of the meaning of a patient's constructs, and it is perhaps worth noting that the overlap scores for LIKE SELF/UNLIKE SELF sorts were among the highest obtained. This type of assumption may result partly from a tendency to forget about the specific task the subject is being given, and an over-readiness to generalize about construct relations from a limited sample of constructs used in a limited context. The task involved is one of describing *people* in terms of certain labels If these labels represented construct dimensions used by a subject to distinguish *between people* (e.g. tall/short, heavy/light) then little or no overlap might appear. However, since it seems likely that all the dimensions involved are used, if at all, to distinguish various types of actions or characteristics *within people*, consistent converse relations between opposites need not be expected.

(2) Kelly indicated that one important feature of the grid method was that the data it provides can be subjected to various analyses, irrespective of the elements or constructs involved, to explore facets of the structural or mathematical features of the construct system being measured. Other workers have used a variety of structural measures (Bannister, 1960; Bieri, 1961). Although these writers are no doubt aware of the limitations of these measures, a newcomer to this field might readily gain the impression that such structural analyses are somehow more secure, more generalizable, and more independent of the specific test situation than are scores derived from the type of content analyses mentioned earlier. The results in 3(*b*), using various forms of intensity scores, indicate that, at least for one such measure, the use of contrasting poles of constructs may provide very different results. Again the present results imply caution in making generalizations from construct data gained in a limited test situation.

(3) In a grid where people are the elements, it may be worth noting that degrees of positive or negative association between constructs for any individual reflect something of the discriminative usefulness of the constructs in *distinguishing between people*. This may not be the main use the person himself makes of his constructs, but this is one piece of information that emerges for the tester. Thus he may make broad anticipations about some of the labels the subject may apply to someone he

classified as, for example, GENEROUS. It may be that many of the people described as GENEROUS are also described as MEAN. In this instance (as for subject 9), the dimension GENEROUS/MEAN, though perhaps meaningful for the subject, does not serve as a useful discriminative dimension for separating out this person's anticipations about people as wholes. Matching element choices for GENEROUS with those for other labels may indicate some other word which is applied more consistently to those people not called GENEROUS; for example, STUPID. The 'peculiar' dimension, GENEROUS/STUPID, might then serve as a potentially more useful discriminator between people, even though not used directly by the subject himself, than either the GENEROUS/MEAN or INTELLIGENT/STUPID dimensions.

This point may have some importance since a number of methods of plotting grid data graphically have been suggested (Bannister, 1965; Slater, 1965). These methods seem to rely on the appropriateness of the contrast poles elicited from or supplied to the subject as end-points of the orthogonal dimensions plotted. If, as appears here, the contrast poles of many constructs may themselves provide essentially 'orthogonal' choices of elements, then many misleading pictorial representations of construct relations may have appeared. Functional contrasts could, however, be extracted in the way suggested above if the personal contrasts given by the subject do not serve as genuine discriminators between the particular type of element involved.

Implications for measurement and research

The present results call attention to the importance in grid measurement of using personally meaningful construct dimensions, and illustrate that when verbal labels are supplied to subjects they may each use different dimensions involving a variety of contrast poles. Experimenters who supply construct labels to subjects as bases for grid sorts should explicitly recognize the range of individual interpretations of these labels by eliciting the personal contrasts involved for each subject. The results also suggest that grid sorts should often be made in relation to *both* poles of each dimension. The grid method is already a fairly time-consuming method of measurement, and this suggestion, if followed, would clearly add considerably to the time involved. This disadvantage may be offset by certain advantages. Considerably more information may be gained about that section of the person's construct system under investigation, and this information may prevent erroneous inferences.

One further suggestion deriving from the present results relates to the graphical plotting of data. When grid scores are being plotted, 'functional' contrasts may sometimes be more useful than personal contrasts for highlighting the between-element discriminations of which the subject is capable.

In relation to broader problems concerning the functioning of constructs and construct systems the present findings may open up some new avenues of exploration. Repeated application of both ends of a set of constructs to any person (element) may provide the basis of a measure of ambivalence, conflict, or potential change of construction. Constructs with considerable overlaps from contrast pole sorts may, however, be those used for purposes other than inter-person discrimination, and differences between people in terms of overlap in any set of constructs may lead in to useful study of differential functions served by different constructs for one indi-

vidual or similar constructs for different individuals. Differences in intensity of matchings at opposite poles may have important bearings on how a person defines his role. Thus within the limits of the constructs involved in this study, subject 1 seemed to have a clearer conceptual picture of SELF than of NOT SELF, while subject 10 appeared to have a more marked set of construct links defining what was UNLIKE SELF rather than what was LIKE SELF. In therapy, or any attempt at construct change, a useful function might be served in elaborating the linkages at one end of a construct dimension rather than at another, and if there were few or no conceptual links with, say, the implicit pole of an important construct, then the patient might be unwilling or unable to change his view of himself on that dimension since change to the opposite pole could precipitate a sense of meaninglessness or anxiety.

Although the idea of the dichotomous use of bipolar constructs may prove useful, the notion of bipolarity may prove more difficult to translate adequately into objective methodology than has yet been suggested. People may readily give apparently clear-cut verbal opposites, but they may sometimes fail to use the verbalized contrasts consistently. Sometimes the actual nature of the contrasts involved may not be clear to the person using the construct, and the tentative opposites he suggests may ill define the vagueness of the fluctuating contrast implied. Again, it seems possible, and even probable, that people, when describing others, do not use specific oppositionality as suggested by the use of single verbal labels, but may rather use amalgamations of ideas in some form of generalized contrast. Furthermore, a person may often use a variety of contrasts to any particular label or general description, and thus his selection of elements may sometimes be controlled by one contrast and sometimes by another.

In addition to all these points, contrasts used in real life almost always involve specific contexts. An employer may be concerned with A's EFFICIENCY at work, and not at making cocktails; or with B's ENERGY in completing the jobs allotted to him, rather than in playing tennis in the evenings. Thus it may be that the construction of a person as KIND in certain contexts has different predictive implications from the construction of him as KIND in other situations. For example, if a person uses KIND as applying to the treatment of animals rather than people, the resulting grid sort might relate very differently to that for INTELLIGENT, than would the sort based on KIND when used to subsume helpfulness to people. This leads in to the idea of the possible use and importance of investigating more of the internal structure of any generalized construct that a person uses than is at present advocated. Thus it may sometimes be important, and could certainly be instructive from the point of view of advancing understanding of the use of constructs and developing more suitable methodologies, to investigate the ramifications of one construct being used by a person. By so doing it would be possible to explore the range of contexts and elements to which each pole could meaningfully be applied, and the implications that any particular use of the construct could involve.

REFERENCES

BANNISTER, D. (1960). Conceptual structure in thought-disordered schizophrenics. *J. ment. Sci.* **106**, 1230–49.

BANNISTER, D. (1962). The nature and measurement of schizophrenic thought disorder. *J. ment. Sci.* **108**, 825–42.

BANNISTER, D. (1965). The rationale and clinical relevance of repertory grid technique. *Br. J. Psychiat.* **111**, 977–82.

BIERI, J. (1961). Complexity-simplicity as a personality variable in cognitive and preferential behavior. In *Functions of Varied Experience*, ed. D. W. Fiske & S. Maddi. Homewood, Ill: Dorsey.

BRIDGES, D. V. (1965). The application of repertory grid technique to schizophrenic thought-disorder. (Unpublished B.A. thesis, University of Reading Library.)

KELLY, G. A. (1955). *The Psychology of Personal Constructs*, vol. 1. New York: Norton.

MAIR, J. M. M. (1966). Prediction of grid scores. *Br. J. Psychol.* **57**, 187–92.

RESNICK, J. & LANDFIELD, A. W. (1961). The oppositional nature of dichotomous constructs. *Psychol. Rec.* **11**, 47–55.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

SLATER, P. (1965). The use of the repertory grid technique in the individual case. *Br. J. Psychiat.* **111**, 965–75.

(*Manuscript received* 2 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 271–282
Printed in Great Britain

SOME PROBLEMS IN REPERTORY GRID MEASUREMENT

II. THE USE OF WHOLE-FIGURE CONSTRUCTS

By J. M. M. MAIR

*Academic Department of Psychiatry, Middlesex Hospital Medical School, London, W.*1

Whole-figure constructs (e.g. LIKE SELF, LIKE FATHER, etc.) have often been used in repertory grid testing to gain indirect assessments of a person's conceptual picture of important figures in his life. Some problems in the use and interpretation of such constructs were explored experimentally. Ten male medical student volunteers acted as subjects. The results indicate that direct ratings of figures and comparable grid scores are not equivalent, and reveal various kinds of discrepancy from expectations concerning: (*a*) relations between one whole-figure construct (LIKE SELF) and personal constructs elicited to define 'like' and 'unlike self' characteristics; (*b*) relations between supplied constructs matching highly with LIKE SELF and elicited constructs defining self and non-self; (*c*) relations among and between elicited 'like' and 'unlike self' constructs. Limitations of the present grid form are discussed in the light of these findings, and interpretations of some of the discrepancies are suggested.

A construct was described by Kelly (1955) as a bipolar dimension subsuming a limited number of events in a person's experience, and used by him as a basis for summarizing and anticipating certain events. A construct is not adequately expressed by describing the elements subsumed by one pole or the other, but by the inclusion of both and the implied contrast between them. Each person is assumed to possess many such dimensions which are organized in a personal hierarchical order in which some are more important than others.

In this system, any person *as perceived* by himself or another is seen as the point of intersection of many personal construct dimensions, in much the same way that the north pole, for example, is defined as the point of intersection of a set of geographical constructions. Kelly translated this idea into concrete form in his repertory grid method; the person being tested is encouraged to provide personally useful constructs, and then to use these to distinguish between a number of elements (e.g. people). Thus each element can be described in terms of its placement on the constructs involved, and something of the person's organization of constructs for dealing with these elements can be inferred from regularities between the various sequences of elements chosen in grid sorts. Some idea of the person's conception of himself or another person can be gained in two main ways from this type of data. The person's own name and/or the names of the other figures of interest can be included as elements, and the constructs applied to these elements can be noted, or alternatively, the name or names of the figures under investigation can be used repeatedly (whether or not the names are included in the element list) as one or more of the figures in the triads from which constructs are elicited.

Both of these methods seem open to conscious modification and influence by the person being tested. Moreover, in recent years a number of studies have appeared using adjectives supplied by the experimenter rather than constructs elicited from the subjects, so the second approach is not always possible, and since some studies

now involve standard sets of elements (such as photographs of unknown people), the first approach has sometimes also been ruled out.

Some workers have sought to gain an indirect measure of some aspects of the conceptual meaning of the self, or any other figure, by using that person as a construct in the grid (whether or not he is included in the element list). Thus the subject is asked to select or rank elements in terms of their similarity in some important respects to the figure in question.

Such *whole-figure constructs* have now been used in a number of investigations (Mair, 1964; Salmon, 1964; Fransella, 1965; Knowles & Purves, 1965; Fransella & Adams, 1966; Fransella & Bannister, 1967; Mair & Boyd, 1967) and in many unpublished clinical studies. They continue to be used because they seem to provide a relatively simple means by which an investigator can gain some idea of the conceptual meaning of a figure in a person's life, while the subject may be unaware of the information he is imparting. He may readily be encouraged to participate since the method makes no demands on him to reveal in detail the reasons for his element choices in the grid sorts. This continuing use of whole-figure constructs seemed to warrant investigation because of a number of inherent theoretical and practical problems in the method.

On theoretical grounds the use of a conglomerate term, such as LIKE SELF or LIKE FATHER, may not be considered strictly to be a construct, since no dimension of similarity and difference is specified. Rather, the subject is free to choose elements on many different construct dimensions, and each dimension could involve a different contrast.

More practically, some unpublished data collected as part of another study (Mair & Boyd, 1967) suggested no great similarity between direct ratings of SELF, FATHER and MOTHER on eleven adjectival scales which were supplied, and grid scores gained by matching grid sorts for LIKE SELF, LIKE FATHER and LIKE MOTHER with sorts based on these same adjectives. In a group of twenty delinquent boys, the average correlations between the direct ratings and the grid scores for SELF, FATHER and MOTHER were 0·45, 0·45 and 0·47 for a split-half grid form, and 0·47, 0·44 and 0·62 for a rank order form. There was, however, a very wide scatter of correlations for different individuals, ranging from −0·62 to 0·92 for the split-half form, and from −0·58 to 0·89 for the rank order grid. It was clear that whatever the rating and grid methods were measuring, the two forms were not necessarily equivalent or even similar for many people. This of itself need not cast doubt on the usefulness of either method, but certainly suggests that a person's direct awareness or assessment of his own or another person's major characteristics may differ widely from those apparently emphasized by the more indirect grid method.

A further consideration aroused concern about the adequacy of using whole-figure constructs in grid measurement. In the usual grid forms it seems difficult, if not impossible, to gain other than an internally consistent and simple (though not necessarily accurate) picture of a person's view of some figure. This consistency seems to be an artifact, partly attributable to the usual rules of grid administration and partly to certain inherent features of the methodology, rather than necessarily a feature of the subject's actual view of the figure under study. The grid method allows measurement of one type of relation between constructs, used in a specific

type of context. Construct relation scores are of the correlational type and are deduced from the degree of similarity in element selection or arrangement resulting from a person's sorts using two or more constructs. Construct dimensions are assumed to be bipolar, and in most applications grid sorts are made only on the basis of one pole of each dimension. Thus there is no opportunity for a person to attribute opposing characteristics to the same person. If this possibility were admitted by requiring grid sorts using both poles of each construct dimension, one might gain a positive matching between, for example, LIKE SELF and both poles of a particular construct (e.g. INTELLIGENT and STUPID). In this way the misleading implication of converse relations implied in the correlational type of score may be partly overcome. This and other problems concerning construct bipolarity have been discussed elsewhere (Mair, 1967).

Even if this aspect of greater complexity were admitted by allowing grid sorts with both poles of constructs, Kelly's type of grid method would still seem to prevent the appearance of certain unique figure constructions. If, for example, a person sees intelligent people as generally energetic, but sees himself as an exception to this rule, being intelligent but lethargic, this type of relation could not show up in grid scores derived from matching a LIKE SELF sort with those for INTELLIGENT and ENERGETIC. Since most of the same people would have been chosen as INTELLIGENT and ENERGETIC, a positive matching between both of these and LIKE SELF would probably appear if the person's decisions on the figure construct were governed largely by the notion of 'intelligence' or some other closely related construct. The result would then suggest that he saw himself as *both* intelligent and energetic. If, however, he tended to choose people as like himself on the basis of lethargy or any mixture of intelligence and lethargy, or any other important construct, a wide variety of construct relation scores between LIKE SELF and these two constructs might appear, and none need reflect the aspect of personal uniqueness noted above.

In attributing meaning to matchings between grid sorts for whole-figure constructs and those for specific construct dimensions, a number of assumptions seem to be involved. The argument seems to run as follows: in grid sorts where elements are chosen because they are like some figure (e.g. self/father/doctor/experimenter) each element may be chosen for a different reason (i.e. on a different construct dimension) and each element left unchosen may be so placed for a different reason; few if any of the elements may have been chosen on the basis of any construct specifically used in the grid and this might be particularly relevant when constructs are supplied rather than elicited. It is assumed, however, that all constructs used by an individual must be interrelated in some way within the complexities of his construct system, and because of this the constructs on which the various elements were chosen as LIKE SELF, for example, will tend to be positively related to each other, and each must at some point in the system be related positively to constructs in the grid which prove to match positively with LIKE SELF, and negatively to those which show a negative matching with LIKE SELF.

Although these assumptions seemed inadequate when stated bluntly, they seemed worth examining in practice, and constituted the present point of departure. The aims were: (*a*) to determine the degree of association between direct ratings and grid scores for SELF, IDEAL SELF, FATHER and MOTHER on eight adjectives; (*b*) to examine the

grid matching scores between sorts for LIKE SELF and UNLIKE SELF and the constructs actually used by each subject to define self and not-self; (*c*) to examine relationships between *supplied* constructs matching highly with LIKE SELF and UNLIKE SELF, and the *elicited* constructs defining self and not-self; (*d*) to examine something of the complexity of 'like self' and 'unlike self' constructions, and the grid's capacity to reflect such complexity, by (i) checking the grid matchings among and between elicited like self and unlike self constructs, and (ii) noting the variety of relations between constructs which show positive or negative matchings with LIKE SELF.

METHOD

Subjects. The same ten male medical student volunteers used previously (Mair, 1967) were involved.

Measures: Grid 1. A split-half grid form using as elements the names of twenty people known well by the subject. Grid sorts were made in terms of four supplied adjectives (KIND, INTELLIGENT, GENEROUS and ENERGETIC) and their opposites (as used by each subject). In addition, sorts were made for MOST LIKE SELF, LEAST LIKE SELF, LIKE I'D LIKE TO BE, LIKE FATHER, LIKE MOTHER and KNOW BEST.

Grid 2. The same twenty elements were used in a split-half format. Grid sorts were made on twelve personal constructs (different for each individual). Six of these dimensions involved some aspects of each person's assessment of characteristics he himself possessed, and six dealt with characteristics the person used specifically to define contrasts to his view of himself. Grid sorts were again made for MOST LIKE SELF.

Direct ratings. Each subject rated SELF, IDEAL SELF, FATHER and MOTHER on 7-point scales for each of the four supplied adjectives used in Grid 1, and their opposites (eight scales each). The extreme and mid-points of the scales were named—Extremely, Average, Not at all.

Procedure. Subjects completed Grid 1 first, and immediately after this the Direct Ratings were obtained. At an average time of 3 days after completing Grid 1, each subject was shown the ten names he had chosen as MOST LIKE SELF, and asked to say what characteristics in each individual underlay their choice. The same procedure was repeated for the ten elements chosen as LEAST LIKE SELF.

For each subject, the experimenter selected from the descriptions defining others' perceived likeness to self, six apparently different characteristics. These were used as the emergent poles of constructs for sorts in Grid 2. Six features defining the antithesis of self were similarly chosen from the explanations for LEAST LIKE SELF choices, and these also served as constructs in Grid 2.

Grid administration and scoring. The general administrative and scoring procedures for split-half grids have been described by Bannister (1960), and the actual procedures followed in the administration of Grid 1 have been presented elsewhere (Mair, 1967). In a split-half grid with twenty elements there are eleven possible matching scores between constructs, going by steps of 2 from 0 to 20. Throughout this study, matching scores of 14 or more are taken to indicate *positive* relations between constructs, scores of 6 or less are taken to indicate *negative* relations, and scores between 8 and 12 as indicative of essentially *chance* association.

RESULTS

In each of the following sections the general expectations concerning results, deriving either from previous findings (1) or from the argument outlined above (2, 3 and 4), are outlined prior to the descriptions of the calculations involved.

(1) *Relations between direct ratings and grid scores*

Expectation. Average correlations between direct ratings and grid scores will be positive, but relatively low.

Spearman rank order correlations were calculated for each subject between the

direct ratings for SELF on eight adjectives (the four supplied adjectives used in Grid 1, and their elicited opposites) and the Grid 1 matching scores derived from comparing element choices in the LIKE SELF sort with those for the same eight dimensions. This comparison between ratings and grid scores was repeated for LIKE I'D LIKE TO BE (Ideal Self), LIKE FATHER, and LIKE MOTHER. The averages for these correlations for the total group (using z transformation) are shown in Table 1.

Table 1. *Group averages and ranges of correlations between direct ratings and grid scores for SELF, IDEAL SELF, FATHER and MOTHER*

	SELF	IDEAL SELF	FATHER	MOTHER
Mean	0·55	0·88	0·62	0·67
Range	−0·29–0·84	0·73–0·93	0·28–0·84	0·25–0·89

(2) *Matchings between MOST LIKE SELF and LEAST LIKE SELF and elicited constructs*

Expectations. The majority of matchings between elicited like-self constructs and the whole-figure construct LIKE SELF will be positive, as will the majority of matchings between elicited unlike-self constructs and UNLIKE SELF. Some may be at a chance level, but none will be negative.

In Grid 1 subjects made grid sorts for MOST LIKE SELF and LEAST LIKE SELF. In Grid 2 they made sorts among the same elements on six constructs specified by them as dimensions of similarity between themselves and others, six specifying dimensions of dissimilarity, and again sorts were made for MOST LIKE SELF.

The matchings between MOST LIKE SELF on Grid 1 and Grid 2 ranged from 16 to 20. There was 92% agreement overall in element choices on these grid sorts.

For each individual, matching scores between LIKE SELF on Grids 1 and 2 and the elicited like and unlike self constructs were calculated; also between the UNLIKE SELF sort on Grid 1 and the elicited constructs. Table 2 shows the percentages of positive, chance and negative matchings of the two LIKE SELF sorts and the UNLIKE SELF sort with the elicited like and unlike self dimensions.

Table 2. *Percentages of positive, chance and negative matchings of two LIKE SELF sorts and an UNLIKE SELF sort, and the elicited like and unlike self constructs*

	Elicited like self constructs			Elicited unlike self constructs		
	Positive	Chance	Negative	Positive	Chance	Negative
Grid 1. LIKE SELF	42	50	8	3	65	32
Grid 2. LIKE SELF	47	48	5	2	58	40
Grid 1. UNLIKE SELF	7	58	35	40	58	2

The percentage figures in each section (positive/chance/negative) were based on 60 scores.

(3) *Relations between supplied constructs matching highly with LIKE SELF and UNLIKE SELF (Grid 1) and elicited constructs defining self and unlike self constructions (Grid 2)*

Expectations. The majority of matchings between supplied constructs relating highly (positively or negatively) with LIKE SELF and/or UNLIKE SELF and elicited constructs defining self or unlike self constructions will be high and in the same

direction. Some of these matchings may be at a chance level, but none will be negative.

The sorts based on the twelve elicited constructs in Grid 2 were matched with those for the eight constructs in Grid 1 (four supplied adjectives and their opposites) giving ninety-six scores for each subject. For each subject the number of positive, chance and negative matchings of each of the six elicited like self constructs with the eight Grid 1 constructs were calculated. The same procedure was repeated for the six elicited unlike self constructs matching with the same eight Grid 1 constructs. The Grid 1 matchings of the eight constructs with MOST LIKE SELF and LEAST LIKE SELF were examined, and those constructs with positive, chance and negative matchings with either of these noted. The overall percentages of positive, chance and negative matchings of these Grid 1 constructs matching positively, negatively or at a chance level with MOST LIKE SELF and LEAST LIKE SELF with the elicited like and unlike self constructs were then calculated. These results are shown in Table 3.

Table 3. *Percentages of positive, chance and negative matchings for eight Grid 1 constructs relating positively, negatively or at a chance level with* MOST LIKE SELF *and* LEAST LIKE SELF *and elicited like and unlike self constructs in Grid 2*

A. Grid 1 construct matchings with MOST LIKE SELF	Grid 1 constructs matching with elicited (Grid 2) like self constructs		
	Positive	Chance	Negative
Positive (n = 96)	46	46	8
Chance (n = 282)	19	66	15
Negative (n = 102)	9	49	42

B. Grid 1 construct matchings with LEAST LIKE SELF	Grid 1 constructs matching with elicited (Grid 2) unlike self constructs		
	Positive	Chance	Negative
Positive (n = 102)	49	46	5
Chance (n = 276)	18	62	20
Negative (n = 102)	9	57	34

The n values refer to the numbers of matchings involved in each set (positive/chance/negative) of percentages.

Table 4. *Percentages of positive, chance and negative matchings among elicited like self and unlike self constructs and between like and unlike self constructs*

	Positive	Chance	Negative
Matchings among like self constructs (n = 150)	32	56	12
Matchings among unlike self constructs (n = 150)	39	49	12
Matchings between like and unlike self constructs (n = 360)	16	54	30

4(*a*) *Relations among and between elicited like self and unlike self constructs*

Expectations. The majority of matchings among elicited like self constructs and among elicited unlike self constructs will be positive, while the majority of matchings between elicited like and unlike self constructs will be negative. Within each of these areas some matchings may be at a chance level, but none will be negative.

For each individual the fifteen matchings among the six elicited like self constructs were calculated, as were the fifteen among the six unlike self constructs, and the thirty-six matchings between the like and unlike self constructs. The group percentages of positive, chance and negative matching scores within each of these three sections were calculated and are shown in Table 4.

4(*b*) *Matchings among elicited constructs which relate positively or negatively to* MOST LIKE SELF

Expectations. The majority of matchings among elicited constructs which relate positively to MOST LIKE SELF will be positive, as will those among elicited constructs relating negatively to MOST LIKE SELF. Some may be at a chance level, but none can be negative.

The number of positive and chance matchings among elicited constructs matching positively with MOST LIKE SELF in Grid 2 was calculated. The same was done for constructs matching negatively with MOST LIKE SELF. From thirty-nine matchings among constructs with positive relations with MOST LIKE SELF, 64% were positive and 36% were at a chance level. In thirty-seven matchings among constructs matching negatively with MOST LIKE SELF, 86% were positive and 14% were at a chance level. It is mathematically impossible for any negative matchings to appear in this situation.

Discussion

In this study an attempt has been made to explore some of the problems underlying the use and interpretation of whole-figure constructs in grid measures. Attention has been focused particularly on one whole-figure construct which has been used widely—LIKE SELF/UNLIKE SELF—and the results presented are illustrative rather than definitive. Problems in ascribing meaning to grid matchings with whole-figure constructs were explored by examining the association between direct assessments of figures on a number of supplied dimensions and comparable grid measures; by examining the congruence in grid terms between whole-figure sorts and sorts based on various elicited personal construct dimensions; by examining the links between matchings of supplied constructs with a figure construct, and the elicited dimensions defining aspects of that figure. They imply limitations of the grid form that was used in representing some of the complexities of self assessment that are apparent when direct elicitation is used.

Direct ratings and grid scores

The average correlations between ratings and grid scores for SELF, FATHER and MOTHER were positive, but modest, while the range of individual variation in each case was considerable. In this instance the average correlations were all higher than those found previously, but many differences in the populations tested and the grids used might have contributed to this result. The most curious result was the high agreement between ratings and grid scores for the ideal self, or LIKE I'D LIKE TO BE, construct. It seems reasonable to suppose that people tend to have a clearer or more definite picture of what they would like to be than of what they actually are. There is some support for this view in that there was a significant tendency for subjects to have more extreme construct matchings with LIKE I'D LIKE TO BE than

with LIKE SELF (Wilcoxon's $T = 5{\cdot}5$, $n = 10$, $P < 0{\cdot}05$, two-tailed), and more extreme ratings on the ideal than on the actual self scales (Wilcoxon's $T = 0$, $n = 9$, $P < 0{\cdot}01$, two-tailed). Although scores from such extremes need not contribute to greater correlations, the use of opposites might have aided such agreement. The finding of greater similarity between direct ratings and grid scores for the ideal rather than the actual self is of some interest when viewed in relation to three earlier studies in which a LIKE I'D LIKE TO BE construct allowed better predictions or discriminations than LIKE SELF (Salmon, 1964; Knowles & Purves, 1965; Fransella & Bannister, 1967). The grid is often described as an indirect measure in that the subjects are seldom aware of the kind of information they are providing, and may thus supply information which they could not themselves directly conceptualize. In this instance it seems that subjects were in some way more aware of their conception of what constituted their postulated future ideal self than their present self, and the other studies imply that in some situations understanding of behaviour may be aided by appreciation of a person's anticipations of future possibilities rather than by his summary of past events. Further studies might usefully explore this area, and provide some explanations of results which are at present simply interesting oddities.

Grid relations between LIKE SELF *and* UNLIKE SELF *and elicited self-descriptions, and between supplied constructs and these same self-descriptions*

The second stage of the study involved another comparison between direct assessments of a kind and grid matching scores. In this instance each subject provided a description of what was and was not like himself by describing ways in which other people resembled and contrasted with himself. These descriptions then served as construct labels to allow sorting among the various elements involved. In the first part of the study, where supplied constructs were used, there seemed little basis for attributing greater significance to either ratings or grid scores. Here, the elicited self-descriptions were used as criteria against which to assess the adequacy of the whole-figure—LIKE SELF—construct matchings.

The results obtained (Table 2) did not fully accord with expectations. It seemed that there may be a greater likelihood that a chance matching score will appear between LIKE SELF and constructs specifically selected as indicating like or unlike self characteristics than either positive or negative scores. On the other hand, if constructs have been specifically chosen to indicate aspects of like self assessments, then there is a greater likelihood that LIKE SELF will match positively rather than negatively with such constructs. This also applies to selections for UNLIKE SELF matched with unlike self constructs. However, in both instances some significant negative matchings did appear.

The other main part of the argument outlined earlier led to the expectation that if, say, LIKE SELF were matched with a number of supplied constructs, then there would be a strong tendency for constructs matching positively with LIKE SELF to be related positively to constructs in each subject's own system defining his self construction, and negatively to constructs specifying unlike self constructions. Supplied constructs matching positively with UNLIKE SELF might be expected to relate in a similar fashion to elicited constructs defining unlike self and self perceptions.

The results shown in Table 3 are similar to those in Table 2. The expectations were not fully supported. Again there was a definite tendency for matchings to be in the expected direction rather than in the opposite direction, but chance level matchings tended to be as frequent as or more frequent than the expected positive or negative relations. Again some matchings contrary to the expectations appear, and the actual grid scores themselves provide no clue as to which construct matchings with the whole-figure construct indicate accurate reflexions of some meaningful association.

These results do not encourage the view that it is easy to derive valid information from grids by a simple process of matching sorts for whole figures with those for other construct dimensions. However, in view of the assumption in construct theory that each person's constructs are organized into some personally useful hierarchy, some being more important than others, it would seem reasonable to suggest that the constructs in Grids 1 and 2 which matched highly with LIKE SELF might be relatively more important for the subjects than the others involved in the grids. This idea was investigated by means of two statistical techniques which have been used previously to specify more or less important constructs in grid use.

First, it has been suggested by Bannister (1965) that constructs which account for relatively more variance in matching scores in a grid may be considered as possibly indicating more important dimensions by comparison with the others. In Grid 1, the matching scores between all the eight adjectives were calculated. For each of these construct labels, the deviations from absolute chance (a score of 10) of its matchings with the other seven were summed regardless of sign. The relative sizes of these eight totals for each individual now indicate the amount of variance in matching scores taken up by each construct in its matchings with the others. For each subject, the constructs matching highly with LIKE SELF (6 or less and 14 or more) were noted, and the totals relevant to these were averaged, and compared by means of Wilcoxon's test for matched pairs (Siegel, 1956), with the averaged totals for the constructs matching at a chance level with LIKE SELF. This procedure was repeated for the twelve constructs (6 like self, and 6 unlike self) in Grid 2.

For Grid 1, no significant difference appeared ($T = 19$, $n = 10$, N.S.), while in Grid 2 the constructs matching highly with LIKE SELF showed a significant tendency to have higher matchings with the other constructs in the grid ($T = 2$, $n = 9$, $P = 0{\cdot}01$, two-tailed). Using this statistical criterion of 'importance' therefore, there seemed to be a tendency for the elicited constructs matching highly with the whole-figure construct to be generally more important in the grid than the others. No such tendency appeared with the supplied construct grid. This may have been due to the particular set of constructs involved, but a similar non-significant result also appeared in another study involving supplied constructs (Mair & Boyd, 1967).

The second statistical approach to the investigation of important dimensions was applied only to Grid 2. Kelly's non-parametric factor analytic procedure for extracting the main clusterings of constructs in grids was used, on the assumption that the more important constructs for each subject would tend to show significant loadings on the first factor. The results indicated that 80% of constructs matching highly with LIKE SELF in Grid 2 showed high loadings (14 or more) on the first factor (hence 20% of constructs matching highly with LIKE SELF did not show high loadings on the first factor); 58% of constructs with high loadings on the first factor showed

high matchings with LIKE SELF (hence 42% of constructs with high loadings on the first factor did not match highly with LIKE SELF); and 50% of constructs *not* matching highly with LIKE SELF had high loadings on the first factor.

Thus while the first of these results seemed to support the idea that the constructs matching highly with the whole-figure construct would be those with high loadings on the first factor (and thus perhaps more important for the subjects), the other findings confused the issue by indicating that only about half of the constructs with high loadings on the first factor were highly related to LIKE SELF, and half of the constructs which were not related highly to LIKE SELF still had high loadings on this factor.

Perhaps no very useful conclusion can be drawn from either of these statistical approaches as to the *psychological* importance of constructs which matched differentially with LIKE SELF. A useful function may have been served, however, in highlighting both the criterion problem which besets the use of grids in this study and others, and the limited usefulness of *post hoc* summary methods for elucidating psychological significance. It seems likely that statistical techniques like those used here will solve very few of the psychological problems which arise in attempts to understand the functioning of construct systems. Instead, methodologies will have to be developed which allow more direct assessment of constructs functioning in important ways; further attempts will have to be made to unravel the logic of what goes on when a subject chooses elements on the basis of certain constructs; and grid measures will have to be used along with others in predictive studies where criteria external to these measures can be used to allow assessment of the relative adequacy of each.

Individual complexity and grid measures

The results in Table 4 show the degrees of association among and between elicited like and unlike self constructs. Only 30–39% of matchings were in the anticipated direction, and higher percentages of contradictory matchings appeared here than in either Tables 2 or 3. The finding that only a third of the matchings were in the expected direction may simply indicate the obvious—that people may view themselves in ways which cannot necessarily be accounted for on one conceptual dimension. The grid method can demonstrate certain aspects of this individual complexity, but no single matching of any whole-figure construct could be associated significantly with all relevant constructs. When Kelly's factor analytic method was used on the grids containing the elicited constructs it appeared that although the majority of constructs had high loadings on the first factor (64%), a considerable number remained to provide high loadings on the second factor (20%), with the remainder constituting the third and subsequent factors.

Although this kind of factorial complexity can be estimated in grids, certain other limitations exist. Some problems concerning the assumption of bipolarity have been discussed elsewhere (Mair, 1967). A further limitation of the present grid form in reflecting some other aspects of individual complexity can be demonstrated from the present results. Among both the like self and unlike self constructs, 12% showed significant negative matchings with the other constructs. Thus a person may place himself at the emergent poles of two constructs which in grid terms may show negative matchings with each other. If only grid scores are considered, and information

about self constructions derived simply from matching LIKE SELF with the other constructs, this type of information (i.e. self placed positively on two or more dimensions which are negatively related in general, or vice versa) could not appear. In the grids used at present, if two constructs match highly (6 or less and 14 or more) with a third construct, then these two constructs will either match positively or at a chance level with each other. They *cannot* show a negative relationship with each other. In the present grid form, two constructs each providing matching scores of 14 with a third construct, cannot show a matching of less than 8 with each other. Any increase in matchings of either with the third will increase the necessary matching between them.

Conclusions

The results of this study do not allow any general conclusions about the usefulness of matchings between whole-figure and other constructs, the meaning that can be attached to them, or the situations in which their use is justified. The legitimate conclusions are more limited, since the overall aim of the study was to explore problems in the use of whole-figure constructs, rather than to provide practical solutions to unclear problems.

The general conclusion is that many problems do seem to arise with the use of whole-figure constructs (though these problems may only differ in degree from those arising with the use of more specific constructs), and methodologies are needed which can resolve some of the difficulties. A number of lines for further investigation have been suggested. The particular conclusions are as follows:

(1) Direct ratings and grid scores do not provide equivalent results in most instances. Both may be useful, but predictive studies are needed to test this possibility.

(2) Sorts for LIKE SELF were matched with sorts for constructs specifically elicited to define aspects of each subject's self description. When such a whole-figure construct is matched with elicited constructs, positive and negative matching scores are likely to reflect similar self placements on these constructs, but many (even the majority) of the chance matching scores may conceal meaningful positive or negative self constructions. A few positive or negative matchings may appear which run completely counter to the subject's direct self-descriptions.

(3) No single matching of a whole-figure construct with other constructs is likely to relate to all constructs which are relevant to adequate description of the figure in question, since most people function on more than one conceptual dimension.

(4) In the present instance, as with many studies, statistical techniques were of limited help in revealing the personal importance of constructs matching differentially with LIKE SELF.

(5) The grid method used can provide only an internally consistent and simple set of construct relationship scores, which may be factorially complex, but which fail to reflect some of the construct linkages demonstrated by comparing directly elicited with grid data.

REFERENCES

BANNISTER, D. (1960). Conceptual structure in thought disordered schizophrenics. *J. ment. Sci.* **106**, 1230–49.

BANNISTER, D. (1965). The rationale and clinical relevance of repertory grid technique. *Br. J. Psychiat.* **111**, 977–82.

FRANSELLA, F. (1965). The effects of imposed rhythm and certain aspects of personality on the speech of stutterers. (Unpublished Ph.D. thesis, University of London Library.)

FRANSELLA, F. & ADAMS, B. (1966). An illustration of the use of repertory grid technique in a clinical setting. *Br. J. soc. clin. Psychol.* **5**, 51–62.

FRANSELLA, F. & BANNISTER, D. (1967). A validation of repertory grid technique used as a measure of political constructing. *Acta psychol.* **26**, 97–106.

KELLY, G. A. (1955). *The Psychology of Personal Constructs*, vol. 1. New York: Norton.

KNOWLES, J. B. & PURVES, C. (1965). The use of a repertory grid technique to assess the influence of the experimenter–subject relationship on verbal conditioning. *Bull. Br. psychol. Soc.* **18** (59), 23A.

MAIR, J. M. M. (1964). Paired-associate learning in neurotics: an investigation of individual differences and the effects of repeated failure. (Unpublished Ph.D. thesis, University of London Library.)

MAIR, J. M. M. (1967). Some problems in repertory grid measurement. 1. The use of bipolar constructs. *Br. J. Psychol.* **58**, 261–270.

MAIR, J. M. M. & BOYD, P. R. (1967). A comparison of two grid forms. *Br. J. soc. clin. Psychol.* **6**, 220–7.

SALMON, P. (1964). The repertory grid as a predictor of experimental conformity in children. *Bull. Br. psychol. Soc.* **17** (55), 7A–8A.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

(*Manuscript received* 2 *January* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 283–289
Printed in Great Britain

THE RELATION OF VERBAL RATE TO SYLLABLE LENGTH OF VERBAL ACTION IN TWO INTERVIEW SITUATIONS

By JAMES T. WEBB and DAVID M. BRISTER
University of Alabama and
Veterans Administration Hospital, Tuscaloosa, Alabama

Undergraduate students were interviewed in standardized (SI) and non-standardized (NSI) situations and their speaking rates measured in syllables per minute. Analyses of verbal rates supported Goldman-Eisler's (1954*b*) finding of an inverse relation between subjects' verbal rate (VR) and length of action, and between the variance of subjects' VR and length of action, but these relations were observed only in the SI situation. Frequency distributions for VRs were found to be normal, by contrast with the positively skewed and leptokurtic distributions she reported. Homogeneity of VR variance across subjects was found in the SI situation, but not in NSI. Thus, the type of interview situation, i.e. the experimenter's behaviour, apparently influences subjects' VR in such a way that the more consistent the experimenter's interview behaviour, the less the variability of the subject's behaviour. Clinical implications are pointed out.

In a naturalistic observation study of language behaviour in diagnostic interviews by three psychiatrists with ten patients, Goldman-Eisler (1954*a*) found mean verbal rate within one dyad (two-person interaction) to be the most consistent temporal aspect of subjects' behaviour in relation to dimensions investigated. In addition, it was found to be the most sensitive aspect of subjects' verbal behaviour when the interlocutor was varied.

In a later study using a similar interview situation with eight subjects Goldman-Eisler (1954*b*) observed the relation of verbal rate to syllable length of actions. An action was defined as the subject's speech between the last word of the interlocutor's preceding utterance and the first word of his subsequent utterance. An analysis of variance indicated that verbal rate discriminated significantly between subjects ($P < 0{\cdot}001$) over the entire range of lengths of actions. Furthermore, as the number of syllables comprising an action increased, verbal rate decreased in a linear fashion for all subjects; hence subjects who were relatively fast talkers in short actions remained so in long ones, despite the generally decreased speech rate found to accompany increases in syllable length of action. A similar finding obtained in the relation between syllable length of action and variance of verbal rate. Variance decreased in an approximately linear fashion for all subjects up to about 100 syllables, at which point the curve apparently became asymptotic. Frequency distributions of scores were found to be positively skewed and leptokurtic for all subjects, with the range extending from 40 to 600 syllables/min. The positive skewness was related primarily to verbal actions of 40 syllables or less which were apparently spoken at a faster rate than actions of more than 40 syllables.

In her earlier study, Goldman-Eisler (1954*a*) did not specify what aspects of the interlocutors' behaviour accounted for changes in subjects' verbal rates across interlocutors. Later she inferred (Goldman-Eisler, 1954*b*) that the variation in the amount of fluctuation across subjects arose from individual difference in flexibility. The relatively constant verbal rate maintained by individual subjects within a given

conversation prompted her to conclude that a general level of stimulation was aroused during particular conversations. She further observed that high verbal rates by a subject either preceded or followed high rates by the interlocutor. In addition, rates increased when one subject interjected a remark into a 'breathing or thinking' pause of the other. Increases apparently due to conversational interruption were short-lived, however, for the rate quickly returned to its former level.

The present investigation replicated and extended portions of Goldman-Eisler's (1954*a*, *b*) studies in an experimental rather than a naturalistic observation setting. Two types of interview situations were used. The first, a standardized interview, was developed by Chapple (1939) and Chapple & Arensberg (1940), and was subsequently refined by Matarazzo, Saslow & Matarazzo (1956) as an experimental approach to the study of non-content speech interactions. This standardization technique allows systematic variation of specific aspects of an experimenter's speech. Many non-content speech variables investigated within this paradigm have been shown to be significant and controlling dimensions of interpersonal interaction almost regardless of content (Matarazzo, 1962; Saslow & Matarazzo, 1959). Verbal rate has not been investigated before in such a standardized interview.

METHOD

Subjects. Four male and two female university undergraduates from introductory psychology courses were used as subjects. They were between the ages of 18 and 21 yr, had no known history of personal emotional disturbances, and their I.Q.s exceeded 110 as measured by the vocabulary subscale of the Wechsler Adult Intelligence Scale.

Procedure. All interviews were conducted individually in a 15 ft. × 15 ft. sound-deadened room overlooked by a one-way vision mirror behind which the observer sat. Two microphones were placed on the table, before experimenter and subject respectively, and the interview was tape-recorded. Each interview lasted 15 min, with the onset, the half-way point, and the end being signalled to the observer by the experimenter through a prearranged, unobtrusive hand signal.

Prior to the interview, subjects were randomly divided into two groups, A and B, each to receive the control (standardized) and the experimental (non-standardized) portions of the interview, but in a different order. For subjects in group A the first half of the interview was non-standardized. The experimenter asked direct questions, varied the length and speed of his speech, and made no effort to control gestures. During the final half of group A interviews, the experimenter structured his behaviour according to the basic standardization rules of Matarazzo (1962), as follows:

(*a*) The experimenter introduces the standardized portion of the interview by a 5 sec utterance (following his signal to the observer).

(*b*) All interviewing must be 'non-directive'. There may be no direct questioning, no probing, and no depth interviewing. The experimenter can reflect, ask for clarification, ask for more information, or introduce a new, general topic area. In general, the experimenter's comments should be non-challenging and open-ended, and related to the subject's past comments or to some new, general topic.

(*c*) All the experimenter's actions must be verbal only, or verbal and gestural at the same time; i.e. the experimenter may not use head nods and other gestures alone.

(*d*) All the experimenter's utterances must be 5 sec in duration.

(*e*) After the subject finishes a comment or other verbal action, the experimenter must respond in less than 1 sec.

In addition, the experimenter was not permitted to interrupt the subject during the standardized portion of the interview. For subjects in group B, the sequence of non-standardized and standardized portions of the interview was reversed.

The purpose of the experiment was explained to subjects as follows: 'I am getting experience

in operating some new tape-recording equipment, and in interviewing people. What I am interested in is finding out student attitudes toward the world situation at this time and our international policies. I wonder if you could tell me some of the feelings and thoughts that you have about it?' This introductory statement was found to provide a topic general and vague enough for all subjects to be able to talk in a relaxed fashion for the 15 min required.

Interview tapes were transcribed verbatim, including repetitions, false starts, and vague sounds such as 'eh' and 'ah'. Verbal rates were measured in syllables/min, since syllables are more appropriate than words for measuring utterance length (Cotton, 1936; Goldman-Eisler, 1954*b*). *Webster's New Collegiate Dictionary* (1960) was used as the criterion of syllabic composition except for the sounds 'eh' and 'ah' which were counted as syllables.

Speaking times were measured with a stopwatch calibrated in 0·1 sec. Interviews were taped at 7·5 in./sec, but played back at 3·75 in./sec so that time could be measured with greater precision. Anticipation of the onset and offset of speech, necessary for accurate measurement, was achieved with the aid of the typescript and by visible marks on tape deck and reel. Measurements were made to 0·1 sec and repeated until agreement was obtained between two consecutive timings. It was assumed that measurement errors were randomly distributed. The average error of measurement (total interview time less the sum of the individual measurements, divided by the number of the individual measurements) was found to be less than 0·03 sec.

Results

For comparison with Goldman-Eisler's (1954*b*) data, verbal actions for each subject were separated into seven categories of syllable length, and mean and median verbal rates were computed for each. In the standardized interview (SI) condition subjects' mean and median verbal rate decreased in an approximately J-shaped fashion as the syllable length of action increased. In the non-standardized interview (NSI) there was no group trend condition and individual differences were striking, particularly in actions of more than 100 syllables in length.

Table 1. *Homogeneity of the variances of different subjects' verbal rates for all categories of syllable length combined, and separately for categories 0–59 and 60+ (Cochran's test)*

	Categories of syllable length	C-value	D.F.*	*P*
Standardized interview	All combined	0·288	6, 19	N.S.
	0–59	0·260	6, 12	N.S.
	60+	0·301	6, 13	N.S.
Non-standardized interview	All combined	0·345	6, 25	< 0·01
	0–59	0·435	5, 10	< 0·05
	60+	0·315	6, 16	< 0·05

* The first D.F. is the number of variances being compared; the second is the largest value of $(n-1)$ for the variances being compared.

Cochran's test for homogeneity of variance (Winer, 1962) indicated that variances across subjects were significantly different ($P < 0·01$) in the NSI condition, but not in the SI condition (Table 1). It was therefore acceptable to plot a composite curve of mean verbal rate during SI (Fig. 1). In spite of the lack of homogeneity, a composite curve of mean verbal rate during NSI was plotted for comparison. In contrast to the J-shaped curve in the SI condition, the NSI curve shows a random relation between syllable length and mean verbal rate. The levelling of the verbal rate curve which Goldman-Eisler (1954*b*) noted at syllable lengths beyond 100 syllables was also evident, particularly in the SI condition.

Table 2. *Subjects' verbal rates*

Subject	Standardized interview			Non-standardized interview		
	Mean	S.D.	*n*	Mean	S.D.	*n*
1	271	59	19	204	43	7
2	266	49	13	324	40	11
3	246	42	11	241	41	26
4	211	68	20	212	66	12
5	261	36	18	232	32	14
6	291	74	17	271	46	12

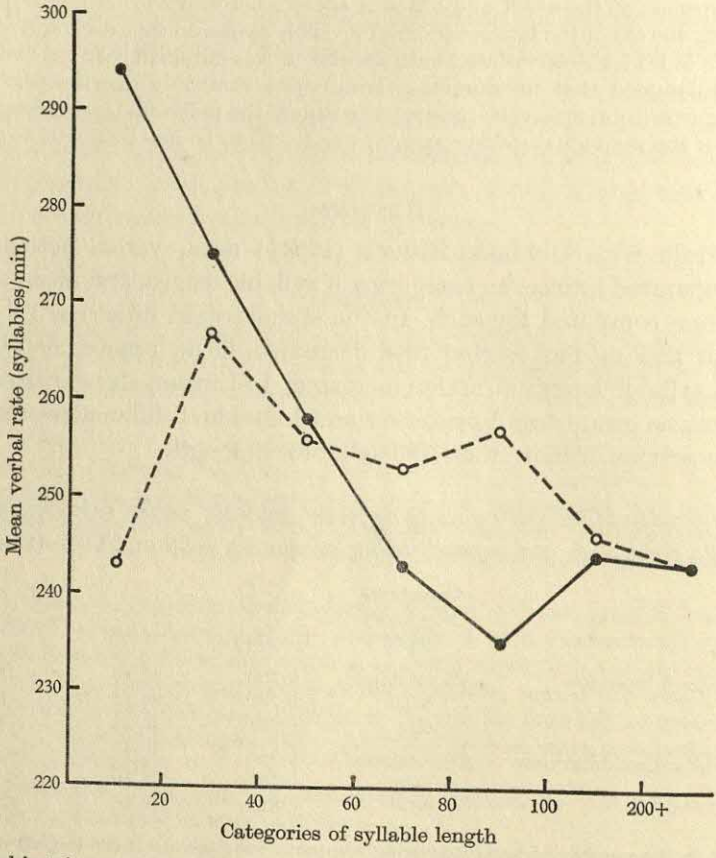


Fig. 1. Subjects' mean verbal rate for actions of different syllable length during standardized interviews (○- - -○) and non-standardized interviews (●—●).

Within-subject variance of verbal rate was also related to syllable length of actions, as Goldman-Eisler (1954*b*) found. Once again, the relation was clearer in the SI condition than in the NSI condition. Individual subjects' S.D. were plotted for each of the two categories (< 60 and $\geqslant 60$ syllables), for the SI and NSI conditions (Fig. 2). In the SI condition S.D. for all subjects decreased in the $\geqslant 60$ syllables category. Since increases or decreases in S.D. are equally likely in a random distribution, the probability of obtaining six decreases out of six by chance alone is 1/64 ($P < 0{\cdot}02$). In the NSI condition only five subjects' S.D. were plotted, since

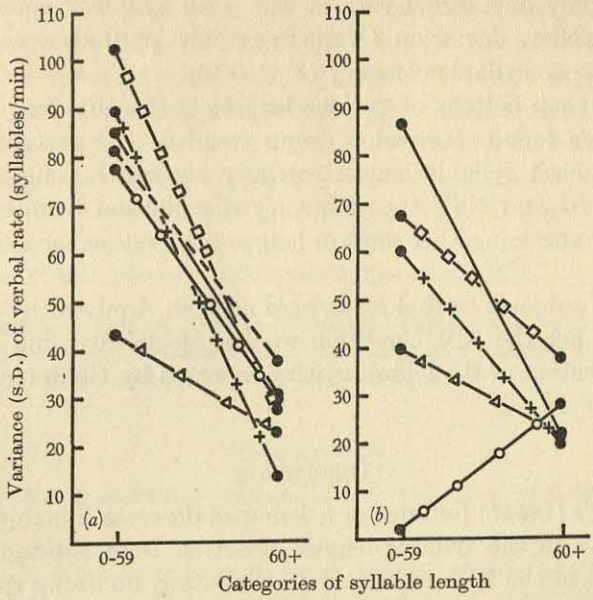


Fig. 2. Variance of verbal rate (S.D.) within individual subjects for actions of different syllable lengths (*a*) during standardized interviews and (*b*) during non-standardized interviews. Each line connects the points for one subject and the same line has been used for the same subject in (*a*) and in (*b*); results could be plotted for only five of the six subjects in (*b*) (see text).

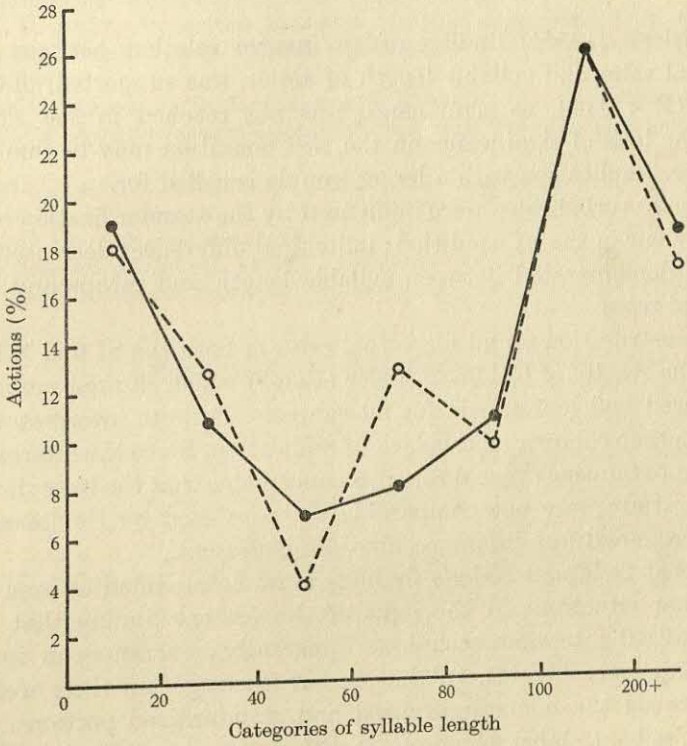


Fig. 3. Percentages of subjects' actions in each category of syllable length during standardized interviews (○- - -○) and non-standardized interviews (●—●).

one subject had only one verbal rate in the < 60 syllables category, preventing calculation of a standard deviation. Of the five results plotted, however, four showed a decrease in the ⩾ 60 syllable category ($P < 0{\cdot}16$).

Subjects did not use actions of various lengths with equal frequency, nor was a normal distribution found. Instead a group trend of the percentage of subjects' actions falling in each syllable length category showed a rough U-shaped curve (Fig. 3) for both SI and NSI, the frequency distribution being smoother for SI. Subjects tended to speak in either short or long actions rather than actions of medium length.

Distributions of subjects' verbal rates were normal. Analyses by χ^2 indicated that in neither the SI nor the NSI condition were subjects' distributions skewed, and there was no indication of the leptokurticity reported by Goldman-Eisler (1954*b*).

DISCUSSION

Goldman-Eisler's (1954*b*) findings of a J-shaped decrease in subjects' mean verbal rate with increases in the syllable length of action were substantiated in the SI condition, but not in the NSI condition. In the latter, no group trend was evident, indicating that a different functional relation between verbal rate and syllable length developed in the NSI as opposed to the SI condition. This finding was interpreted as indicating that the experimenter's behaviour in the NSI condition influenced subjects' speech rates. Furthermore, individual differences apparently became more pronounced.

Goldman-Eisler's (1954*b*) finding of an inverse relation between variance of subjects' verbal rates and syllable length of action was supported, but only in the SI condition ($P < 0{\cdot}02$), as significance was not reached in the NSI condition ($P < 0{\cdot}16$). The lack of significance in the NSI condition may be due to the small sample number. Replication with a larger sample is called for.

Hence subjects' verbal rates were influenced by the standardization of the experimenter's behaviour in the SI condition; individual differences decreased, and strong relations were demonstrated between syllable length and means and variances of subjects' verbal rates.

The normal distribution found for verbal rates in both the SI and NSI conditions contradicted the results of Goldman-Eisler (1954*b*) which showed verbal rate to be positively skewed and leptokurtic for all subjects. Definite interpretations cannot be made at this time concerning this lack of replication. Since the interview situation has been shown to influence verbal rate it is conceivable that the interviews employed in the present study were not comparable to those used by Goldman-Eisler. The possibility of cross-cultural differences likewise is present.

It appears that Goldman-Eisler's findings must be modified as applying only to certain interview situations in the light of the present finding that Matarazzo's (1962) standardized interview technique apparently generalizes in its controlling effect to subjects' speech rates. In the present investigation there were numerous differences between the non-standardized and standardized portions, and at this point it is impossible to label which of the differences were critical. Further investigation is needed to determine which aspects of the standardized interview exert

control over a subject's verbal rate. But it can be stated that the more consistent the experimenter's interview behaviour, the less is the variability of subjects' verbal rates, both across and within subjects, while variations in the experimenter's behaviour within and across subjects are strongly reflected in subjects' behaviour.

The strong possibility exists, therefore, that an experimenter's behaviour in therapeutic or diagnostic interviews influences a subject's non-verbal speech measures. Hence the clinical use of non-content speech measures in diagnostic or mental status evaluations should be more cautious in the light of the present findings, and the findings of Saslow & Matarazzo (1959) and Matarazzo (1962).

The authors extend their appreciation to Dr Michael Dinoff for his valuable methodological and stylistic suggestions, and to Messrs Jerry Cardwell, Jim Foulks, Warren Simpson, and Robert Valachovic for their assistance in analysing the data.

References

Chapple, E. D. (1939). Quantitative analysis of the interaction of individuals. *Proc. natn. Acad. Sci. U.S.A.* **25**, 58–67.

Chapple, E. D. & Arensberg, C. M. (1940). Measuring human relations: An introduction to the study of the interaction of individuals. *Genet. Psychol. Monogr.* **22**, 3–147.

Cotton, J. (1936). Syllabic rate: a new concept in the study of speech rate variation. *Speech Monogr.* **3**, 112–17.

Goldman-Eisler, F. (1954*a*). A study of individual differences and of interaction in the behaviour of some aspects of language in interviews. *J. ment. Sci.* **100**, 177–97.

Goldman-Eisler, F. (1954*b*). On the variability of the speed of talking and its relation to the length of utterances in conversation. *Br. J. Psychol.* **45**, 94–107.

Matarazzo, J. D. (1962). Prescribed behaviour therapy: suggestions from clinical research. In *Experimental Foundations of Clinical Psychology*, ed. A. J. Bachrach, pp. 471–509. New York: Basic Books.

Matarazzo, J. D., Saslow, G. & Matarazzo, R. G. (1956). The interaction chronograph as an instrument for objective measurement of interaction patterns during interviews. *J. Psychol.* **41**, 347–67.

Saslow, G. & Matarazzo, J. D. (1959). A technique for studying changes in interview behaviour. In *Research in Psychotherapy*, ed. E. A. Rubenstein & M. B. Parloff, pp. 125–59. Washington, D.C.: American Psychological Association.

Webster's New Collegiate Dictionary. (1960). Springfield, Mass.: Merriam.

Winer, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

(*Manuscript received* 2 *January* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 291–299
Printed in Great Britain

AGE DIFFERENCES IN THE TEMPORAL INTEGRATION OF LANGUAGE

By F. I. M. CRAIK* and P. A. MASANI

Medical Research Council Unit for Research on Occupational Aspects of Ageing, and Department of Psychology, University of Liverpool

Results of a previous experiment suggested that old subjects are no worse than young subjects at remembering meaningless verbal material but are at a disadvantage when dealing with meaningful material. This result contradicts established views, so two short experiments were carried out to examine age differences in short-term memory for various types of verbal material. The results confirmed the original observation that old subjects are at a greater disadvantage when dealing with highly meaningful material and also showed old subjects to be more vulnerable to the effects of supraspan interference. An explanation of the results is put forward in terms of age differences in temporal integration.

In a previous experiment (Heron & Craik, 1964) examining age differences in learning of meaningful and meaningless material it was noted that there was no age difference in the memory span for meaningless sounds while an age difference was found when meaningful words were used. The meaningless material consisted of the digits 1–9 in Finnish while the equivalent digits in English formed the material for the meaningful lists. After familiarization with the sound of Finnish digits, subjects were given a standard 'digit span forward' test with auditory presentation and immediate verbal recall. The result of this part of the experiment was that while both old and young subjects had the same span for Finnish digits, the young group had the higher mean English digit span. When the difference between English and Finnish spans was taken as a score for each subject, a highly significant age difference ($P < 0{\cdot}001$) was found between the two sets of scores.

It was concluded from this result that older subjects were no worse at remembering completely meaningless sounds in relation to younger subjects, but were at a disadvantage when the material was in some way meaningful. It was tentatively suggested that the well-documented age decrements in memory (Welford, 1958, 1964) were due more to the decreasing efficiency of some coding process than to storage difficulties as such. That is, memory for meaningless material was held to reflect some fairly pure retention process, while meaningful material was considered to be more amenable to recoding and 'chunking'. It was suggested that this latter process is perhaps the locus of the memory difficulties of older subjects.

This viewpoint is opposed, however, by the conclusions from previous research (Ruch, 1934; Gilbert, 1941; Korchin & Basowitz, 1957). These studies suggest that older subjects have greater difficulty in learning and remembering unfamiliar, meaningless material. Korchin & Basowitz, for example, concluded that 'older subjects...are proportionately more deficient in the learning of materials in which the facilitative effects of prior experience are minimized'. Similarly, Kay (1959) suggested that 'with familiar material older subjects may not show anything like

* Now at Birkbeck College, University of London.

the same deficit as they do with unfamiliar'. The general point of view held by previous workers thus seems to be that older subjects are at less of a disadvantage in memory and learning tasks when the material is familiar and meaningful, as in this case they can fill in parts of predictable sequences from their past experience. With unfamiliar, meaningless material their past experience is not relevant, and may even be negative in its effect, and therefore much larger age differences should be apparent when this type of material is used.

Two short experiments were carried out to explore this difference in conclusions. The terms 'familiarity' and 'meaningfulness' are used rather interchangeably here to refer basically to the ease with which the material can be recoded. Verbal material was used in both experiments and 'recodability' was manipulated by varying the sequential dependencies between words. In the first experiment a digit span technique was used to determine the maximum amount of different types of material which could be remembered, while the second experiment measured recall for lists of words beyond the span.

Experiment I

The short-term memory span for three types of verbal material was obtained for subjects ranging in age from 20 to 79 yr. The first type of material, English text sentences, was considered to be highly familiar, meaningful and amenable to recoding procedures. The second type of material, colour names, was not so amenable to recoding. For the third type of material, common proverbs with their word order scrambled were used; this was considered to be an example of material which would be disruptive to normal coding procedures. It was predicted that an age decrement would be found for the first and third types of material but not for the second. That is, old subjects would be at a disadvantage when the material was amenable to coding (text) and when inappropriate recoding was possible (scrambled proverbs), but no age difference would be found with material which merely had to be stored and recalled in relatively uncoded form (colour names).

Method

Procedure

The subject was told to repeat back each list of words exactly as he heard them. The experimenter read each list (the text sentences at a slow conversational speed and the other lists at 1 word/sec), the subject then recalled each list immediately after presentation. A digit span technique was used, the number of items in successive lists being increased each time until the subject failed to recall the list correctly. Two lists were given at each length and the score was taken to be the longest message length at which at least one list was correctly recalled, taking items and correct order into account. The lists were presented in the same order to each subject: text, colour names and finally scrambled proverbs.

Subjects

Five men and five women represented each of the six decade age-groups, 20's to 70's. Each group was matched on verbal intelligence as this is a measure known to decline little with age. Mill Hill Vocabulary scores (Set B Synonyms) were 36·0 for all groups except the 50's and 60's; mean scores for these two groups were both 36·1.

Results

The mean length of list correctly recalled is shown for each type of list and each age-group in Table 1. Trend analyses of variance (Brownlee, 1960) were carried out to assess age differences in the length of list recalled. An age decrement was found in ability to recall text ($F = 8{\cdot}26$; D.F. 1, 58; $P < 0{\cdot}01$) and scrambled proverbs ($F = 8{\cdot}08$; D.F. 1, 58; $P < 0{\cdot}01$), but was not found in ability to recall colour names ($F = 0{\cdot}50$; D.F. 1, 58; N.S.).

Table 1. *Memory span for three types of verbal material obtained from subjects aged* 20–79 *yr.*

	Age-groups					
Material	20–29	30–39	40–49	50–59	60–69	70–79
Text sentences	20·5	19·8	20·8	20·0	18·5	17·2
Colour names	6·2	6·1	6·6	6·2	6·0	5·9
Animal names	7·0	7·0	6·9	6·4	6·4	6·3

Discussion

These results are in agreement with the present authors' prediction from the results of Heron & Craik (1964). That is, where the possibility of transforming or 'chunking' verbal material is at a minimum (colour names), old subjects can remember as many items as younger subjects; when, however, this possibility is present (text lists), young subjects can make more use of it. Thus while subjects of all ages can remember longer lists of text than colour names, the older subjects are less able to make use of the structure and redundancy in meaningful verbal material. The age decrement in retention of scrambled proverbs might be interpreted as showing the use of inappropriate coding procedures by the older subjects; they still group words in their most likely order. Alternatively this decrement might be looked on as demonstrating the greater interfering effect on the older subjects of intrusion of the correct order of the proverbs.

Experiment II

A second exploratory experiment on age differences in the retention of verbal material was carried out on the same subjects. In this case meaningfulness was again manipulated by varying the degree of contextual constraint between words. The word lists of different orders of statistical approximation to English generated by Miller & Selfridge (1950) were used as material. The lists chosen were those of 10 and 30 words at zero, first, third and fifth orders of approximation to English, as well as lists of standard text. The predictability of words in these lists, owing to increasing semantic and syntactic constraint, increases as the lists go from zero order to text.

The prediction on the basis of Heron & Craik would be that age differences should be least with the zero-order lists (randomly chosen words with very low transitional probabilities) and greatest with the text lists. On the basis of the previous research cited above, on the other hand, it would be predicted that age differences would be least for recall of the meaningful text lists.

Method

Procedure

The 10-word and 30-word lists generated by Miller & Selfridge (1950) were recorded on tape at 1 word/sec. In each case the five lists were recorded in the same order (0, 1, 3, 5 and text). The subject was instructed to listen to each list as it was played back and then to write down as many of the words as he could, in any order. No time limit on recall was imposed. All subjects were given the list in the same order (10-word lists zero order to text, then 30-word lists zero order to text) and the score calculated was the number of words correctly recalled regardless of order.

Subjects

The subjects were the same as those in Expt I. Thus the ages 20–79 were represented and age-groups were matched on verbal intelligence.

Table 2. *Mean numbers of words recalled by subjects of three age-groups from word lists of different statistical approximation to English* (*Young*, 20–39 *yr.*; *Middle*, 40–59 *yr.*; *Old*, 60–79 *yr.*)

		Degree of approximation to English				
	Age-group	0	1	3	5	Text
10-word lists	Young	4·6	5·6	8·3	9·8	9·9
	Middle	4·0	5·2	8·3	9·8	9·8
	Old	3·6	3·9	6·4	9·5	9·7
30-word lists	Young	5·8	8·9	13·5	15·5	20·1
	Middle	5·1	8·1	11·5	13·3	18·5
	Old	3·1	5·9	9·0	10·4	13·9

Table 3. *Results of trend analyses of variance on recall scores of young, middle-aged and old subjects for* 30-*word lists of different degrees of statistical approximation to English*

	F	D.F.	*P*
Difference of slope from zero			
Young	148·1	1, 98	< 0·001
Middle	184·8	1, 98	< 0·001
Old	121·9	1, 98	< 0·001
Difference between slopes			
Young–Middle	0·78	1, 196	N.S.
Middle–Old	3·02	1, 196	N.S.
Young–Old	5·92	1, 196	< 0·05

Results

For ease of presentation the results were grouped into three large age-groups, each of twenty subjects: Young (20s + 30s), Middle (40s + 50s), and Old (60s + 70s). The mean numbers of words recalled for each age-group and experimental condition are shown in Table 2. This table shows that a ceiling effect was in operation for the 10-word lists in that virtually all subjects correctly recalled the fifth order and text lists. The 10-word lists were therefore not included in further analyses, except in one particular case considered below. Trend analyses of variance were carried out on the 30-word list data and the significance of the slopes (zero order to text) is given in Table 3. Table 2 also shows that the Young group's recall scores on the 30-word lists increased

markedly between zero order and text, while the Middle and Old groups' scores increased to a lesser extent. The significance of differences between these rates of improvement was also assessed by the trend analyses.

Table 3 shows that all three age-groups improved their recall scores very significantly from zero order to text. The difference in slopes between the Young and Middle groups is not significant, the Middle–Old difference approaches significance and the Young–Old difference is significant beyond $P < 0{\cdot}05$. Thus it can be concluded that while all subjects recall more words as the lists become more meaningful, the younger subjects were able to make more use of the contextual constraints inherent in the higher order approximations to English. Older subjects are thus at a relative disadvantage when dealing with highly structured, meaningful material, a conclusion which is in accordance with the present authors' hypothesis.

Although the difference in slopes between the Young and Old groups shown in Table 3 reaches significance, the differences between the means given in Table 2 are

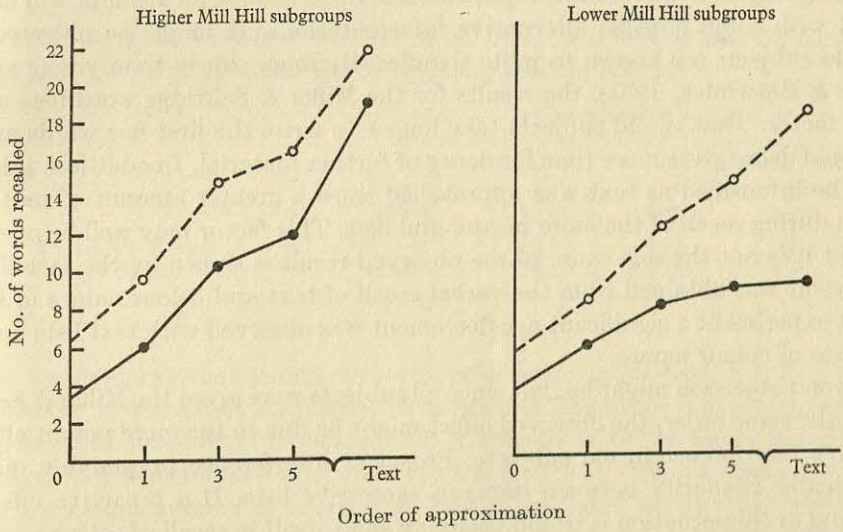


Fig. 1. Mean number of words recalled from 30-word lists of different degrees of statistical approximation to English by young (○ - - - ○) and old (●—●) subjects divided into higher and lower Mill Hill subgroups.

not very great. It seemed possible that the absence of marked differences was due to the subjects tested being of extremely high verbal intelligence. Accordingly the Young and Old groups were split into two subgroups, each of ten subjects: those of very high verbal intelligence (mean Mill Hill Vocabulary = 38·5; groups referred to as Young–high and Old–high); and those of moderately high verbal intelligence (mean Mill Hill Vocabulary = 33·6; groups referred to as Young–low and Old–low). It should be pointed out that the 'low' groups are only relatively low; they still represent the 75th percentile of Mill Hill Vocabulary scores while the 'high' groups represent the 95th percentile. The mean scores of these subgroups are shown in Fig. 1. This figure shows that the Young–low group's scores are at a slightly lower level than those of the Young–high group, but there is no difference in their rates of improve-

ment as text is approached. This rate of improvement is at least maintained by the Old–high group, although again their scores are at a lower absolute level. Subjects in the Old–low group, however, are much less able to make use of the increasing contextual constraints in the higher order lists. Thus there is no age difference in ability to make use of contextual constraint in subjects of very high verbal intelligence (Young–Old difference between slopes, $F = 0{\cdot}03$; D.F. 1, 96; N.S.), while there is a marked age difference in this ability in subjects of moderately high verbal intelligence (Young–Old difference between slopes, $F = 14{\cdot}56$; D.F. 1, 96; $P < 0{\cdot}01$). This analysis shows that when more typical subjects are considered, older subjects are at more of a relative disadvantage when dealing with meaningful material than the results shown in Tables 2 and 3 indicate. On the other hand this relative disadvantage is not shown by subjects of very high verbal intelligence.

Discussion

Before going on to discuss the implications of these results, an attempt will be made to deal with some possible alternative interpretations. It might be objected that, since old subjects are known to write significantly more slowly than young subjects (Birren & Botwinick, 1951), the results for the Miller & Selfridge word lists are due to this factor. That is, old subjects take longer to write the first few words and this additional delay gives more time for decay of further material. In addition, this effect would be intensified as text was approached since a greater amount of material is written during recall of the more meaningful lists. This factor may well be operating, but that it is not the sole cause of the observed result is shown by the fact that the same result was obtained from the verbal recall of text and colour names in Expt I. In that experiment a significant age decrement was observed with text lists but none with lists of colour names.

A second objection might be that since all subjects were given the Miller & Selfridge lists in the same order, the observed effect might be due to the more potent effects of proactive interference in old subjects. Proactive interference presumably increases with greater similarity between items in successive lists. If a proactive effect was operating in this situation it would therefore show itself in recall of colour names. As no age decrement was found for that material it is concluded that an age difference in proactive interference was not a relevant factor in these experiments.

A third objection to the present writers' interpretation is that if ratios between young and old subjects' scores are taken rather than arithmetic differences, the age differences on the Miller & Selfridge lists look somewhat different. When the Old group's score is expressed as a percentage of the Young group's score the results are as follows:

List ...	0	1	3	5	Text
(O/Y) %	53	67	67	67	69

Looked at in this way, old subjects appear to be at the greatest disadvantage when dealing with the relatively meaningless unrelated word lists. This conclusion is strengthened by the fact that there is a fairly large age difference with the zero order lists while the authors' point of view should predict little or none.

The explanation suggested to account for this feature of the results is that basically

old subjects are indeed able to remember as many unrelated words as the young subjects (as shown by their retention of colour names in Expt I) but their score is depressed by their greater liability to interference when the word span is exceeded. This effect of 'supraspan interference' is probably a special case of inter-item interference as discussed by Melton (1963) and others. Melton suggested that the 'chunk' was the relevant unit of interference, so that the more chunks there are in a list, the greater is the effect of inter-item interference. In the case of lists longer than the span, perception and integration of further chunks would interfere with the retention of those chunks in the span. The possibility that old subjects are more prone to supraspan interference has previously been suggested by Welford (1958) and by Kay (1959).

If this factor is indeed operating in the present case it follows that there should be much less of an age difference in recall of unrelated words with the 10-word lists, as in that case there is less supraspan material. Table 4 shows that this is so; there is a smaller age difference for the zero order 10-word list both in absolute scores and also with differences expressed as values of t.

Table 4. *Mean number of words recalled from* 10-*word and* 30-*word lists of unrelated words by young and old subjects*

		Words recalled			
		10-word lists		30-word lists	
	n	Mean	S.D.	Mean	S.D.
Young (20–39 yr.)	20	4·65	1·15	5·85	1·62
Old (60–79 yr.)	20	3·60	1·11	3·10	1·04
t		2·84		6·14	

Table 4 also shows that young subjects recall more words from the 30-word list than from the 10-word list; hence another factor working in the opposite direction to supraspan interference must be postulated. It appears to be the factor described by Murdock (1960) who found that, whatever the length of list, his student subjects recalled a constant number of words from lists of unrelated words plus a percentage of the remainder. On the basis of the present results it appears that while young subjects can recall more unrelated words as list length increases, this tendency is outweighed for old subjects by the negative effects of supraspan interference.

It is therefore concluded that for lists of unrelated words the age difference is caused, or at least greatly amplified, by supraspan interference. As the Miller & Selfridge lists approach text the number of items per chunk increases, subjects can remember more items and there are fewer 'supraspan chunks' to cause interference. The age difference in words recalled is greatest at the text end, however, so another factor must be operating. This factor is judged here to be the older subjects' inability to make as full use of the structure and redundancy inherent in spoken English.

Discussion and conclusions

At first sight it is difficult to reconcile the present results with those of such previous workers as Ruch (1934), Gilbert (1941) and Korchin & Basowitz (1957). Differences in the material and in the experimental procedure may, however, account for the bulk

of the discrepancies; for example, Korchin & Basowitz employed a paced schedule during recall, giving their subjects 4 sec in which to read the stimulus, recall and report the correct association. It is certainly plausible, as Korchin & Basowitz themselves point out, that the effects of this pacing were more severe on the older subjects and especially with the more difficult meaningless material. Similarly, since virtually all the material used in these earlier studies exceeded the immediate memory span, it may be that supraspan interference was particularly potent for the unfamiliar material used.

The two main conclusions which are drawn from the present experiments are, first, that old subjects are at a greater disadvantage when dealing with highly meaningful verbal material, and secondly, that they are more vulnerable to the effects of supraspan interference.

The first conclusion is interpreted here as showing an age decrement in temporal integration. That is, words in spoken or written discourse are held and integrated over time in the subject's attention or perceptual systems, and are then transferred to memory in suitably encoded 'chunks'. Thus the input systems must simultaneously perceive, hold, integrate and store the verbal material. In the case of unrelated words, digit sequences or colour names, there is little opportunity for integrating and recoding the material into a more economical form and in this case little if any age decrement is observed. When the sequential dependency between words is high, however, as in written or spoken English, there is much more opportunity for integration and recoding; it appears to be in this part of the process that old subjects have difficulty.

This interpretation is in general agreement with the results of Wallace (1956). She found that when parts of visual displays were shown serially, so that subjects had to integrate the material in order to identify the total display, older subjects were nearly as good as young subjects at identifying simple displays but there was a progressive age decrement as the material became more complex. The greater difficulty which older subjects experienced with the complex material was attributed largely to the greater amount of temporal integration required with such material. It is suggested that there is a strong parallel between increasing complexity of display in Wallace's experiments and increasing contextual constraint in the present experiments.

The deficit shown by older subjects in recoding and integration could be attributed to one of two factors; either old subjects have forgotten to some extent the structure of language and so meaningful verbal material carries less redundancy for them, or the difficulty might be due basically to the necessary switching of attention between the operations of perception and integration. Both factors are possible sources of the age decrement although the present writers favour the second possibility. Not only does this explanation agree with the results of previous studies which have shown an age decrement in situations where the subject must divide his attention between aspects of perception and memory (Speakman, 1958; Broadbent & Heron, 1962), but the explanation also applies to the second finding of the present experiments. The observation that old subjects are more vulnerable to the effects of supraspan interference is in accordance with the point of view that short-term retention and perception are mediated in part by the same mechanism, and, while all subjects show an adverse effect on short-term retention if they must deal with incoming material

simultaneously, old subjects show this effect to a greater extent. It seems likely that the involvement of some central organizing or decision channel is necessary for such operations as perception, integration and storage over intervals of more than a few seconds; the suggestion (Welford, 1958) that this central channel may suffer a reduction in its capacity with increasing age would account neatly for most of these findings.

The final general point which may be made about these results is that they point to registration rather than storage as being the locus of age differences in short-term memory.

The authors would like to express their thanks for many useful discussions with their colleagues and for technical assistance received, both in the M.R.C. Unit and the Department of Psychology, University of Liverpool.

References

Birren, J. E. & Botwinick, J. (1951). The relation of writing speed to age and to the senile psychoses. *J. consult. Psychol.* **16**, 399–405.

Broadbent, D. E. & Heron, A. (1962). Effects of a subsidiary task on performance involving immediate memory by younger and older men. *Br. J. Psychol.* **53**, 189–98.

Brownlee, K. A. (1960). *Industrial Experimentation.* London: H.M.S.O.

Gilbert, J. C. (1941). Memory loss in senescence. *J. abnorm. soc. Psychol.* **36**, 73–86.

Hearnshaw, L. S. (1956). Temporal integration and behaviour. *Bull. Br. psychol. Soc.* **28** (30), 1–20.

Heron, A. & Craik, F. I. M. (1964). Age differences in cumulative learning of meaningful and meaningless material. *Scand. J. Psychol.* **5**, 209–17.

Kay, H. (1959). Theories of learning and aging. In *Handbook of Aging and the Individual*, ed. J. E. Birren. Chicago: University of Chicago Press.

Korchin, S. J. & Basowitz, H. (1957). Age differences in verbal learning. *J. abnorm. soc. Psychol.* **54**, 64–9.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.* **2**, 1–21.

Miller, G. A. & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *Am. J. Psychol.* **63**, 176–85.

Murdock, B. B. (1960). The immediate retention of unrelated words. *J. exp. Psychol.* **60**, 222–34.

Ruch, F. L. (1934). The differentiative effects of age upon human learning. *J. gen. Psychol.* **11**, 261–86.

Speakman, D. (1958). Cited by A. T. Welford, *Ageing and Human Skill*, p. 251. London: Oxford University Press.

Wallace, J. G. (1956). Some studies of perception in relation to age. *Br. J. Psychol.* **47**, 283–97.

Welford, A. T. (1958). *Ageing and Human Skill.* London: Oxford University Press.

Welford, A. T. (1964). Experimental psychology in the study of ageing. *Br. med. Bull.* **20**, no. 1, 65–69.

(*Manuscript received* 18 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 301–313
Printed in Great Britain

WATER, FOG AND THE SIZE–DISTANCE INVARIANCE HYPOTHESIS

By HELEN E. ROSS
Department of Psychology, University of Hull

Objects generally appear too large under water. The usual explanation for this is that they are magnified by the change of refractive index between air and water. However, they are only magnified because they are optically nearer. If size constancy operates normally they should be seen the correct size but nearer, or enlarged in size but farther than the optical distance. The problem was investigated by requiring sixteen divers to make size and distance judgements of a range of white disks of different sizes set at various distances. Judgements were made for the same display both under water and on land. It was found that size judgements were greater under water, and that this corresponded to relative overestimation of distance. A weak form of size–distance invariance was found both under water and on land, but the relation between size and distance judgements was not the same in both conditions.

Overestimation of size and distance was also found in a fog on land. It was suggested that the overestimation of distance in a fog and under water was due to the exaggerated reduction of brightness contrast with distance, which increases 'aerial perspective'.

It is commonly observed that objects in water appear larger than in air. Skin-divers are apt to report seeing enormous fish, but if the fish are brought to the surface they acquire more credible proportions. The usual explanation is that objects are magnified by the change of refractive index between water and air in the diver's face-mask. The optical effect (Fig. 1) is to locate the image of the object at approximately three-quarters of its physical distance, and thus to increase the angle subtended at the observer's eye by 4/3. In this calculation it is assumed that the angular size of the object is fairly small; otherwise the optical displacement and magnification are greater, and 'pin-cushion' distortion occurs (Barnard, 1961). It is also assumed that the eye is exactly at the water–air interface; and the refraction due to the glass in the face-mask is ignored.

It is necessary also to consider size constancy. A change in angular size is not a sufficient explanation for a change in apparent size, since objects which are brought nearer on land do not normally appear much larger. If correct geometrical relations between apparent size and apparent distance hold under water, an object should appear the correct size at the correct optical distance, or too large and farther than the optical distance.

The increased apparent size of objects under water suggests that their apparent distance is too far. This was found to be so in an experiment in the swimming baths. Twenty subjects were trained on land to adjust the distance of a white disk to 4 ft. from their noses, by sliding it along a horizontal rod; when asked to set it to the same *apparent* distance under water the mean setting was 4 ft. 5 in. If the optical effect were the only consideration the same apparent distance should be about 5 ft. 4 in. When the subjects were asked to set the disk to the same *physical* distance under water, the mean setting was 4 ft. 3 in.—greater than would be expected if subjects had set it at three-quarters of the same apparent distance. The majority of subjects reduced their settings in the expected direction, but five subjects over-corrected in

the wrong direction. Questioning revealed that this was either through incorrect beliefs about the optical effect, or through faulty reasoning about the direction in which to correct. (These were Cambridge science students!) The moral of the second part of this experiment is that individual variability in distance judgements may be due to different beliefs rather than different perceptions.

The first part of the experiment shows that apparent distance is greater than would be expected from the optical distance. This is probably due to 'aerial perspective' or reduction in brightness contrast. Brightness contrast drops off comparatively rapidly under water (Duntley, 1962), even in the 'clear' water of a swimming-pool. Many experiments have shown that changes in the brightness or colour of an object affect its apparent distance, but it seems likely that the main cue is the brightness

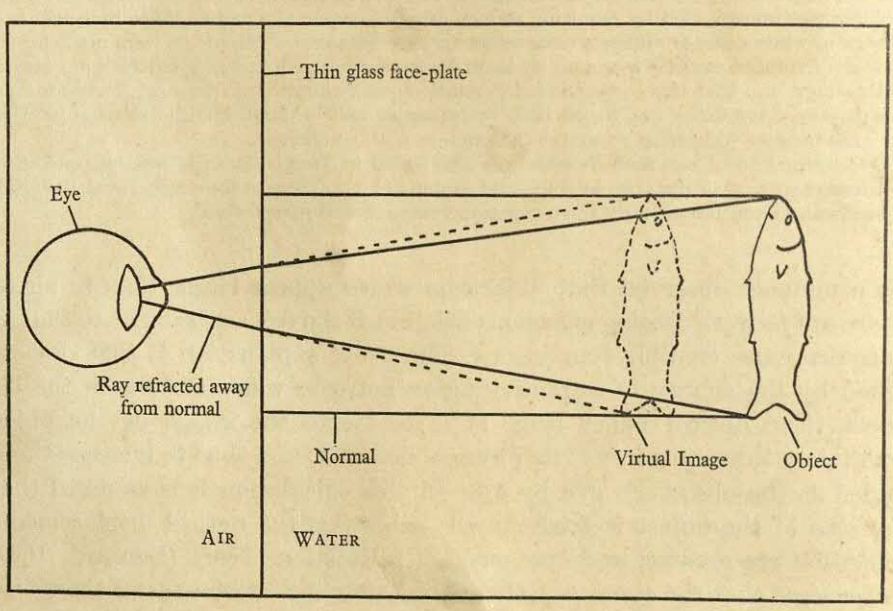
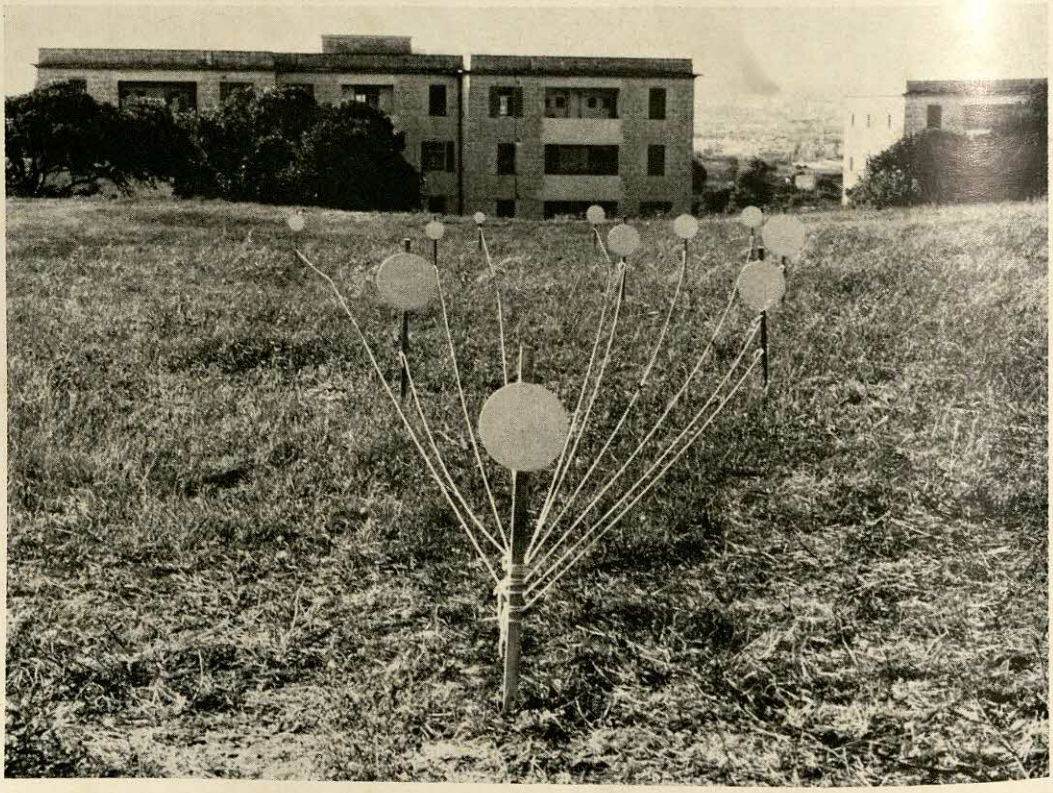


Fig. 1. Rays passing from water into air are refracted away from the normal, since the refractive index of water is 1·33 times that of air. The lens system of the eye (omitted for simplicity) forms a real inverted image on the retina, corresponding to that of an object at about three-quarters of its physical distance from the air–water interface. The angle subtended by the image is thus about 4/3 larger than in air.

contrast between the object and its background (Mount, Case, Sanderson & Brenner, 1956). Changes in brightness contrast are sufficiently effective to upset stereoscopic distance settings (Fry, Bridgman & Ellerbrock, 1949), objects of lower contrast appearing farther away. It was also shown in an earlier paper (Ross, 1965) that changes in the clarity of the water affected size-constancy matches, and this was probably due to changes in the apparent distances of the objects.

Most size-constancy experiments are concerned with size estimates. However, the reasons for changes in constancy ratios cannot be fully understood unless distance estimates are obtained at the same time. This is rarely done, perhaps because of the difficulty of devising convenient methods for measuring apparent distance. The size–distance invariance hypothesis has been questioned (Epstein, Park & Casey, 1961), partly on the grounds that judged size generally increases with distance, while



Top: The range of disks under water at Wied-iz-Zurrieq, Malta.

Bottom: The range of disks on land at Luqa, Malta.

judged distance decreases. Recent studies in which apparent size and distance were measured at the same time generally show that the ratio of apparent size to apparent distance is not invariant over distance (Rump, 1961; Gogel, Wist & Harker, 1963; Gogel, 1964). Some other studies, however, show at least a weak form of size–distance invariance (McDonald & O'Hara, 1964; Hartman, 1964).

Adjustment methods under water are extremely cumbersome and time-consuming. It was decided to forgo all adjustments and rely instead on direct numerical estimates of the true size and distance of objects. The task is similar to magnitude estimates or ratio-scaling techniques, except that the estimates are supposed to refer to physical and not subjective values.

In these experiments objects of different sizes were placed at various distances, and subjects were required to write down a list of their true sizes and distances. The same display was used under water and on land. This made it possible to investigate the relation between size and distance judgements at different distances, and whether the same relations existed under water and on land. These are two distinct questions. If size is incorrectly perceived under water, this could be because distance is incorrectly perceived (relative to the optical distance), or because the ratio of perceived size to perceived distance changes in water.

By using direct estimates rather than matching techniques it is possible to compare perception on land and under water. Comparison by matching would involve rapid alternation between land and water, which would be highly inconvenient. It is assumed, instead, that both land and water estimates are made by comparison with the same internal reference scale. (Such a scale could be based on memory, or the body could be used as a constant yardstick.) This assumption is convenient, but it may not be correct. It is possible that experienced divers change the scale under water, as they learn to adjust to new distance cues. It is also possible that there is a sharp distinction between apparent distance and judged distance which is not revealed by judgements of physical distance. Despite these theoretical difficulties the method gave promising results.

Experiment I: Land and water

Method

The test objects were ten disks, made from 'Dural' and painted white. They varied in diameter from 5·5 to 10 in., in ½ in. steps. An extra disk, 5 in. in diameter, served as the observation post. Each disk was riveted to a 'Dural' stake, so that the bottom of the disk was 15 in. above the ground. The distances were measured by lengths of white nylon cord, which were attached both to the observation post and to the stakes of the test disks. The disks were arranged radially, subtending a physical angle of 20° as measured by a protractor (equivalent to a visual angle of 27° under water). The cords were left in position, as they enabled the diver to identify the disks by counting them from left to right. Coloured beakers were placed between every pair of cords to aid identification. This was important, as the farthest disks were barely visible. The distances and sizes of the disks, from left to right, were: 60 ft. 6·5 in.; 15 ft. 8·5 in.; 45 ft. 6 in.; 75 ft. 5·5 in.; 75 ft. 9·5 in.; 30 ft. 7·5 in.; 45 ft. 8 in.; 60 ft. 9 in.; 30 ft. 10 in.; 15 ft. 7 in. In this arrangement there was one relatively large and one relatively small disk at each of five distances, in order to investigate the effect of size on distance judgements. The range was erected under water on an open sandy sea bed at Wied-iz-Zurrieq, Malta, at a depth of about 90 ft. The range is shown in Plate 1, which demonstrates the loss of contrast with distance. The 'black body visibility'—the farthest distance at which a large black object can be seen (Duntley, 1962)—varied between 75 and 105 ft. Testing generally took place between noon and 5 p.m.

Sixteen divers were used as subjects. Their ages ranged from about 20 to 40 years and their diving experience varied. They were mainly members of university diving expeditions and an air-force diving club.

They were required to kneel at the observation post and write down on a 'Formica' board an estimate of the true size (diameter) and true distance of each disk. A chart with ten columns was drawn on the board, numbered from left to right, corresponding to the spatial arrangement of the disks. Subjects were allowed to use the metric system or yards, feet and inches, whichever they preferred. All subjects used inches for the sizes, and the majority used feet for the distances (yards and metres were sometimes used). Most divers took about 10 min to complete the chart.

When the underwater tests were finished, the range was re-erected on land. A large open field at R.A.F. Luqa was used. The spatial arrangement was exactly the same as under water. The land display is shown in Plate 2. The same subjects made a second series of size and distance judgements on land, a week or two after their underwater test. For practical reasons it was not possible to test any subjects on land first, but it seems highly unlikely that the difference in results was due to the order of testing. The land results were almost identical to those of the control group in Expt II, who were tested only on land.

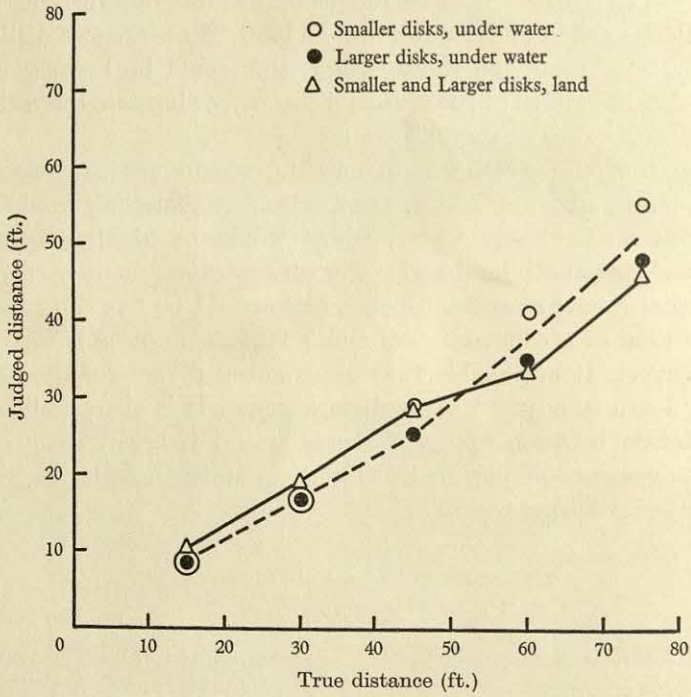


Fig. 2. Mean judged distance as a function of true distance.

Results and discussion

Distance judgements

The mean distance estimates under water and on land are shown in Fig. 2. The mean underwater estimates were always greater than three-quarters of the land estimates. Beyond 30 ft. the underwater judgements were systematically influenced by the relative size of the disks: the smaller of two disks at the same distance was judged farther away. Disk size did not affect distance judgements on land.

An analysis of variance, summarized in Table 1, showed that disk size was a significant variable, though mainly operative under water; however, the interaction between disk size and distance did not reach a significant level. It is probable that the

relative size of the disks was not an effective distance cue on land because it was countermanded by more reliable cues such as the texture gradients on the ground. Subjects were therefore able to show 'distance constancy' despite size change (Smith, 1955; Over, 1961). The ground-texture gradients were much less obvious under water, and the white cords were barely distinguishable from the sandy bottom. It was not clear where the posts of the further disks met the bottom (see Plate 1), so that relative size was one of the few cues available.

Table 1. *Analysis of variance of the distance judgements in Expt I*

Source	S.S.	D.F.	M.S.	F	P
Between locations	128·79	1	128·79	4·63	< 0·05
Between distances	59750·28	4	14937·57	537·32	< 0·001
Between sizes	187·58	1	187·58	6·75	< 0·025
Between subjects	14383·36	15	958·89	34·49	< 0·001
Locations × distances	901·45	4	225·36	8·11	< 0·001
Distances × sizes	210·09	4	52·52	1·89	N.S.
Locations × sizes	226·13	1	226·13	8·13	< 0·01
Locations × subjects	6936·03	15	462·40	16·63	< 0·001
Distances × subjects	10172·22	60	169·54	6·10	< 0·001
Sizes × subjects	354·69	15	23·65	0·85	N.S.
Locations × distances × sizes	205·72	4	51·43	1·85	N.S.
Locations × distances × subjects	5602·86	60	93·38	3·36	< 0·001
Locations × sizes × subjects	357·35	15	23·82	0·86	N.S.
Distances × sizes × subjects	1571·52	60	26·19	0·94	N.S.
Residual	1667·68	60	27·80		
Total	102655·73	319			

The variability between subjects was greater under water than on land, and increased with distance. For example, the estimates for 75 ft. varied between 30 and 66 ft. on land, but between 20 and 135 ft. under water. Not all subjects reacted to water in the same way: the majority overestimated far distances in comparison with their land judgements, but a few underestimated them. It is not clear how far this represents differences in perception or differences in expectation and reasoning.

The distance judgements do not appear to be strictly linear with physical distance. The underwater estimates tended towards increasing overestimation and the land estimates towards increasing underestimation. There was some evidence of non-linearity in the land data ($P < 0{\cdot}05$) when testing the significance of the difference between η^2 and r^2 (Lewis, 1960, p. 124), but not in the underwater data. The non-linearity in the land data was probably due to the dip in the curve at 60 ft., rather than to marked curvilinearity. (The reason for this dip is discussed later.) Several authors have found, using different methods, that true distance on land is increasingly underestimated (Epstein *et al.*, 1961). The reason for the non-linearity is not entirely clear. One possibility is that subjective magnitude is always a power function of physical magnitude (Stevens, 1962). However, when subjects are asked to judge physical rather than subjective magnitude, Stevens's arguments do not seem applicable. An alternative explanation is that distance judgements involve ratio estimates, and that ratio estimates will give rise to power functions (with exponents other than 1·0) if some factor is present which biases the judgements (Restle, 1961, pp. 207 f.). Another possibility is that subjects respond to distance cues which change non-linearly with distance. Accommodation and convergence, for example, may act as cues at near but

not at far distances. Some cues, such as binocular disparity and ground perspective, must be lost sooner under water than on land, being obscured by low brightness contrast. A cue which is exaggerated rather than diminished under water is change of brightness contrast. Fry & Bridgman (1942) have shown that distance judgements increase exponentially as contrast decreases: a given change of contrast is a more effective cue when contrast is low than when it is high. This is probably because contrast changes rapidly over near distances, and more slowly as distance increases (Duntley, 1962): therefore a given change in contrast represents a greater change in physical distance at far distances. This may account for the shape of the underwater distance curve; but further experiments would be necessary to show that it was not due to the method of judgement.

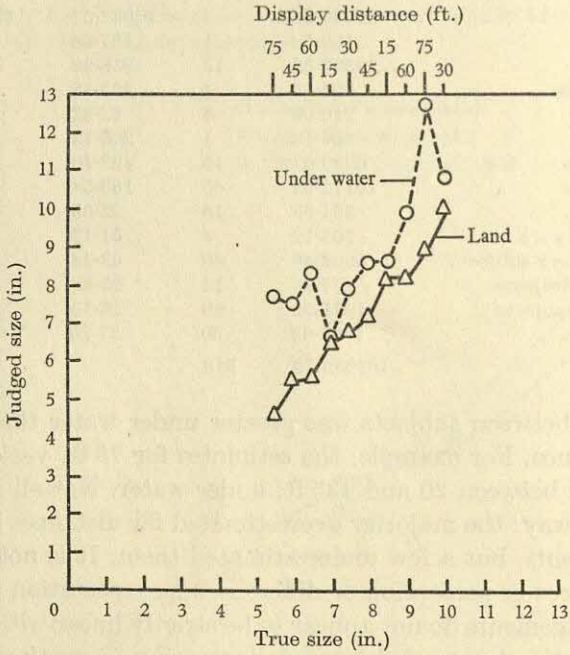


Fig. 3. Mean judged size as a function of true size. The distance at which each size of disk was displayed is indicated at the top of the diagram.

The dip in the distance curve at 60 ft. on land was probably due to the particular geometrical arrangement of the display. The two disks at 75 ft. appeared to be at the centre of the display, while those at 60 ft. appeared to be on the outside of them, creating an effect like the Ponzo illusion (converging lines). Consequently 60 ft. was relatively underestimated and 75 ft. relatively overestimated. In Expt II the geometrical arrangement from left to right was deliberately altered (though the same size–distance combinations were used) and the irregularity disappeared. The irregularity did not occur under water, presumably because the perspective effect was lost in the distance.

Size judgements

The mean size judgements are shown in Fig. 3. As expected, the underwater judgements were considerably larger than the land judgements. An analysis of

variance of the data is summarized in Table 2. As with the distance data, the variability of the judgements was greater under water.

There seem to be two factors upsetting a smooth linear relation between judged and true size. The most obvious factor is the distance of the disks (indicated at the top of Fig. 3). The underwater judgements may appear very erratic, but were systematically affected by distance. At 15 ft. the judged sizes did not differ much between land and water, since the underwater distance estimates were not much greater than three-quarters of the land estimates, while at 75 ft. both sizes and distances were considerably overestimated. Variability of land size judgements is less obvious; but the farther disks were slightly underestimated, both in size and distance. The 7·5 and 9·0 in. disks were the most clearly underestimated, and these were both placed at 60 ft. The size underestimation may be due to the same perspective effect which caused the distance underestimation at 60 ft.

Table 2. *Analysis of variance of the size judgements in Expt I*

Source	S.S.	D.F.	M.S.	*F*	*P*
Between locations	237·18	1	237·18	90·29	< 0·001
Between sizes	778·77	9	86·53	32·94	< 0·001
Between subjects	989·44	15	65·96	25·11	< 0·001
Locations × sizes	103·23	9	11·47	4·37	< 0·001
Locations × subjects	540·43	15	36·03	13·72	< 0·001
Sizes × subjects	410·52	135	3·04	1·16	N.S.
Residual	354·66	135	2·63		
Total	3414·23	319			

Apart from the irregularities due to distance, the size estimates appear to be a linear function of physical size over the range tested. There was no significant departure from linearity in either set of data. This is in agreement with the findings of Stevens & Galanter (1957), Ekman & Junge (1961), Stevens & Guirao (1963) and Teghtsoonian (1965) that magnitude estimates or ratio estimates of length give linear scales (or, equivalently, power functions with exponents of about 1·0). On the other hand, if straight lines were fitted through the data of Fig. 3 they would not go through the origin. Power functions with exponents greater than 1·0 might describe the data more suitably. A possible explanation for this is that, though one-dimensional judgements were required, the objects were two-dimensional. Since area increases as a power function of diameter, judgements of diameter may be distorted by changes in area.

Relation between size and distance judgements

The ratios of judged to true size and judged to true distance are shown in Fig. 4. Optical distance rather than physical distance is used for the underwater ratios. It is evident that strict size–distance invariance does not hold: the size ratios are not the same as the distance ratios and they are not constant over distance. However, a weak kind of invariance holds, in the sense that the size ratios change with the distance ratios over distance. Both ratios decrease with distance on land and increase under water.

The question remains whether the relation between retinal size, judged size and judged distance is the same under water as on land. Can underwater size judgements

be explained purely on the basis of land size constancy and the changed apparent distance under water? If so, the ratio of water to land size judgements should be the same as the ratio of water to three-quarters land distance judgements. These

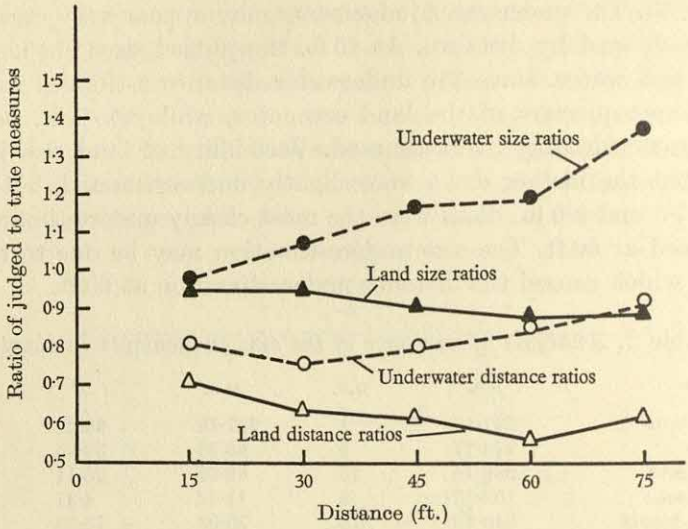


Fig. 4. The ratios of judged to true measures as a function of distance. The underwater distance ratios are based on optical, not physical, distance.

Table 3. *Geometric mean ratios of underwater to land size judgements, and underwater to three-quarters land distance judgements, for each subject*

(Subjects are ranked, as far as possible, according to degree of diving experience in the Mediterranean, with the more experienced divers at the top of the list.)

Subject	Size ratio	Distance ratio
1	1·33	1·50
2	1·24	1·48
3	1·11	1·08
4	0·63	0·74
5	1·62	1·39
6	1·47	1·59
7	1·05	1·38
8	0·91	1·13
9	1·52	1·24
10	1·46	1·49
11	1·57	1·96
12	0·59	1·02
13	1·07	1·12
14	1·19	0·85
15	0·92	1·12
16	2·14	1·58
Geometric mean	1·18	1·25

ratios are similar, as can be deduced from Fig. 4, but they are not identical. The geometric mean of the size ratios was 1·18, and of the distance ratios 1·25. Size was not overestimated under water as much as might be expected from the overestimation of distance. The ratios of individual subjects were far less lawful than has been

suggested by the mean results. Table 3 shows the geometric mean size and distance ratios for each subject. Some subjects showed wide discrepancies between their size and distance ratios, and some subjects *underestimated* size or distance under water. The subjects are arranged in Table 3 in approximate order of diving experience in the Mediterranean. There is some suggestion (though not statistically significant) that inexperienced divers overestimated both size and distance more than experienced divers. The distance judgements of the more experienced divers—many of whom knew the geography of the diving site extremely well—were in fact very similar to their land judgements; their distance ratios are therefore near 1·33, taking the optical effect into account.

There are various possible explanations for the discrepancies between size and distance ratios. Perceptual learning may occur through exploration of the environment (Taylor, 1965); or subjects may deliberately compensate for supposed effects. The average results suggest that a small degree of learning or compensation occurred, since size overestimation was generally not as great as distance overestimation. Individual differences may be due to varying experience or to varying conceptions about the state of nature and the laws of mathematics. Another possibility is that various non-intellectual factors may disturb constancy ratios under water. Nitrogen narcosis might be such a factor. The efficiency of divers using compressed air is noticeably impaired by a depth of 100 ft., due to nitrogen narcosis (Kiessling & Maag, 1962; Baddeley, 1966). Narcosis might directly affect the constancy mechanism, and not merely the intellectual ability to reason about size–distance relations. For example, nitrous oxide makes distant objects appear larger (Steinberg, Legge & Summerfield, 1961) and this may be because it alters constancy ratios rather than increasing apparent distance (Legge, 1965). It is also possible that a reduction of sensory input (Schultz, 1965), or a change of proprioceptive information (Ross, 1965), may alter the ratios.

Experiment II. Fog

Introduction and method

It has sometimes been suggested that overestimation of size under water or in a fog is due to blurring of contours by scattered light. Blurring could increase the apparent size of a bright object, quite apart from any constancy effect. Blurring of contour can also act as a cue to distance (Fry & Bridgman, 1942) and could therefore increase apparent size by constancy, whether the object were brighter or darker than its background. However, it has been shown both theoretically and experimentally that no significant blurring of outline occurs due to fog. The evidence is reviewed by Middleton (1952, pp. 78 f.).

If the argument in the first part of this paper is correct, sizes are overestimated under water largely because reduced contrast leads to the overestimation of distance. One might expect to find similar effects on land, in a fog. There is plenty of anecdotal evidence that this occurs, but I am unaware of any experimental evidence. Experiments are difficult to arrange, since fogs do not occur to order and the visibility does not remain constant for long. During the autumn and winter of 1965–66 there were only about 7 days in which the visibility was reduced to 100 ft. in the city of Kingston-upon-Hull, and the fog generally dissipated by 10 a.m. On one of these

mornings I was able to erect the display on a lawn in the university and test six subjects before the fog cleared. The visibility for large dark objects varied between 100 and 400 ft. during this time. Since the fog was white, the white disks were less visible than black objects, and the visibility of the display was not much greater than in the underwater experiment.

Subjects

The six subjects were unknown students who happened to be passing by, and it was not possible to collect them together again for a control experiment in clear weather. A different group of nineteen psychology students served as controls in brilliant sunshine later that morning.

Apparatus and procedure

The apparatus and procedure were the same as before, except that the spatial arrangement of the disks was rearranged. This was done to break up the perspective effect of the previous display. The new display, from left to right, was: 30 ft. 10 in.; 75 ft. 9·5 in.; 15 ft. 7 in.; 60 ft. 9 in.; 45 ft. 8 in.; 75 ft. 5·5 in.; 30 ft. 8·5 in.; 60 ft. 6·5 in.; 45 ft. 6 in.; 15 ft. 8·5 in. Unfortunately the range was erected hurriedly without a protractor, and the total angle subtended by the outer posts was rather greater than 20°. This may have had some effect upon the results.

Results and discussion

The mean distance estimates of the two groups of subjects are shown in Table 4, and a summary of an analysis of variance of the data in Table 5. The two groups differed significantly, the fog group estimating the distances to be nearly twice as

Table 4. *Mean distance estimates by six subjects in a fog and nineteen subjects in clear weather*

	True distance (ft.)				
	15	30	45	60	75
Clear group					
Larger disks	9·80	18·84	28·53	35·68	42·58
Smaller disks	10·21	19·07	28·13	35·76	45·16
Fog group					
Larger disks	16·63	25·50	58·00	72·00	89·00
Smaller disks	13·67	26·63	57·25	71·00	93·50

Table 5. *Analysis of variance of the distance judgements in Expt II*

Source	S.S.	D.F.	M.S.	*F*	*P*
Between groups	28364·15	1	28364·15	65·95*	< 0·001
Clear group					
Between distances	27178·35	4	6794·59	317·35	< 0·001
Between sizes	558·72	1	558·72	26·10	< 0·001
Between subjects	9570·01	18	531·67	24·83	< 0·001
Residual	3553·39	166	21·41		
Fog group					
Between distances	47795·23	4	11948·81	27·78	< 0·001
Between sizes	3·19	1	3·19	0·01	N.S.
Between subjects	30861·09	5	6172·22	14·35	< 0·001
Residual	21076·30	49	430·13		
Total	168960·43	249			

* Using the residual error term in the fog group, as the larger estimate of residual error.

great as the clear group. The results for the clear group were almost identical with the land results in Malta, except that the irregularity at 60 and 75 ft. was absent. Four of the six fog subjects overestimated the distances, the overestimation tending to increase with distance. However, there was no significant departure from linearity in the data of either group. The size of the disk was not a significant variable in the fog data, perhaps because of the small number of subjects. Contrary to the land data

Table 6. *Mean size estimates by six subjects in a fog and nineteen subjects in clear weather*

	True size (in.)									
	5·5	6·0	6·5	7·0	7·5	8·0	8·5	9·0	9·5	10·0
Distance (ft.)...	75	45	60	15	30	45	15	60	75	30
Clear group	5·39	5·54	5·79	7·16	7·21	7·97	9·03	9·68	9·82	10·09
Fog group	5·50	6·00	7·33	8·00	8·50	11·50	11·17	12·33	11·17	12·00

Table 7. *Analysis of variance of the size judgements in Expt II*

Source	S.S.	D.F.	M.S.	*F*	*P*
Between groups	113·98	1	113·98	72·60*	< 0·001
Clear group					
Between sizes	571·77	9	63·53	72·52	< 0·001
Between subjects	121·25	18	6·74	7·69	< 0·001
Residual	141·93	162	0·88		
Fog group					
Between sizes	358·82	9	39·87	25·37	< 0·001
Between subjects	116·15	5	23·23	14·79	< 0·001
Residual	70·68	45	1·57		
Total	1494·58	249			

* Using the residual error term in the fog group, as the larger estimate of residual error.

Table 8. *Ratios of judged to true measures at different distances for the mean estimates of the fog and clear groups of subjects*

	Distance (ft.)				
Ratios	15	30	45	60	75
Clear distance	0·67	0·63	0·63	0·60	0·59
Clear size	1·04	0·99	0·96	0·98	1·01
Fog distance	1·01	0·87	1·28	1·19	1·22
Fog size	1·23	1·17	1·22	1·25	1·09

in Malta, it was a significant variable with the clear group, the smaller disks being generally judged farther away than the larger disks at the same distance. Perhaps this was because the greater angular separation of the posts made it less obvious that pairs of posts were at the same distance. In addition, the smooth lawn may have provided fewer distance cues than the rough grass on Malta.

The mean size estimates of the two groups are shown in Table 6, and a summary of an analysis of variance of the data in Table 7. The estimates of the clear group were very similar to those of the land experiment in Malta, though slightly larger. As predicted, the estimates of the fog group were significantly larger than those of

the clear group. They were similar to those of the underwater experiment, though the irregularities were not so closely related to changes in distance. There was again no significant departure from linearity. The ratios of judged to true distance and judged to true size are shown for the fog and clear groups in Table 8. The ratios for the clear group show the same general pattern as the land data in Malta, though the decrease with distance is not so marked. The ratios for the fog group are much more erratic, but are similar to the underwater data, suggesting a slight increase with distance.

In conclusion, it seems that the overestimation of size in a fog or under water is due to the relative overestimation of distance. However, lawful relations between size and distance judgements under different conditions are only evident in the average results of fairly large groups of subjects. The laws governing the judgements of individuals are still enshrouded in mist.

The diving experiments were carried out while I was a research student at the Psychological Laboratory, University of Cambridge, supported by the Science Research Council. The fog experiments were carried out at the University of Hull. I am grateful to the E. M. Pratt Musgrave Fund for a travel grant to Malta. I should like to thank Dr A. Baddeley for advice on the experimental method; Mr R. Ward for help in constructing the apparatus; Dr J. N. Lythgoe, Dr C. C. Hemmings and others for much theoretical advice and practical help while diving; Dr J. N. Lythgoe for the underwater photograph, and Dr A. E. Dorey for the land photograph; all the divers from R.A.F. Luqa, Cambridge University and elsewhere who served as subjects; and Mr A. B. Royse for statistical advice.

REFERENCES

BADDELEY, A. D. (1966). The influence of depth on the manual dexterity of free divers—a comparison between open sea and pressure chamber testing. *J. app. Psychol.* **50**, 81–5.

BARNARD, E. E. P. (1961). Visual problems under water. *Proc. R. Soc. Med.* **54**, 755–66.

DUNTLEY, S. Q. (1962). Underwater visibility. In *The Sea*, ed. N. M. Hill, vol. 1, pp. 452–5. New York: Interscience.

EKMAN, G. & JUNGE, K. (1961). Psychophysical relations in visual perception of length, area and volume. *Scand. J. Psychol.* **2**, 1–10.

EPSTEIN, W., PARK, J. & CASEY, A. (1961). The current status of the size–distance hypotheses. *Psychol. Bull.* **58**, 491–514.

FRY, G. A. & BRIDGMAN, C. S. (1942). The influence of atmospheric scattering in range finding. (O.S.R.D. Rep. 930. Office of Publications Bdg., Dept. of Commerce, Washington, D.C.)

FRY, G. A., BRIDGMAN, C. S. & ELLERBROCK, V. J. (1949). The effect of atmospheric scattering on binocular depth perception. *Am. J. Optom.* **26**, 9–15.

GOGEL, W. C. (1964). Visual perception of spatial extent. *J. opt. Soc. Am.* **54**, 411–16.

GOGEL, W. C., WIST, E. R. & HARKER, G. S. (1963). A test of the invariance of the ratio of perceived size to perceived distance. *Am. J. Psychol.* **76**, 537–53.

HARTMAN, A. M. (1964). Effect of reduction on the relationship between apparent size and distance. *Am. J. Psychol.* **77**, 353–66.

KIESSLING, R. J. & MAAG, C. H. (1962). Performance impairment as a function of nitrogen narcosis. *J. appl. Psychol.* **46**, 91–5.

LEGGE, D. (1965). Analysis of visual and proprioceptive components of motor skill by means of a drug. *Br. J. Psychol.* **56**, 243–54.

LEWIS, D. (1960). *Quantitative Methods in Psychology*. New York: McGraw-Hill.

MCDONALD, R. P. & O'HARA, P. T. (1964). Size–distance invariance and perceptual constancy. *Am. J. Psychol.* **77**, 276–80.

MIDDLETON, W. E. K. (1952). *Vision through the Atmosphere*. Toronto: University of Toronto Press.

MOUNT, G. E., CASE, H. W., SANDERSON, J. W. & BRENNER, R. (1956). Distance judgment of colored objects. *J. gen. Psychol.* **55**, 207–14.

OVER, R. (1961). Distance-constancy. *Am. J. Psychol.* **74**, 309–10.

RESTLE, F. (1961). *Psychology of Judgment and Choice*. New York: Wiley.

ROSS, H. E. (1965). The size-constancy of underwater swimmers. *Q. J. exp. Psychol.* **17**, 329–37.

RUMP, E. E. (1961). The relationship between perceived size and perceived distance. *Br. J. Psychol.* **52**, 111–24.

SCHULTZ, D. P. (1965). *Sensory Restriction: Effects on Behavior*. New York: Academic Press.

SMITH, O. W. (1955). Distance constancy. *J. exp. Psychol.* **50**, 374–80.

STEINBERG, H., LEGGE, D. & SUMMERFIELD, A. (1961). Drug-induced changes in visual perception. In *Neuro-Psychopharmacology*, ed. E. Rothlin, vol. 2, pp. 392–6. Amsterdam: Elsevier.

STEVENS, S. S. (1962). The surprising simplicity of sensory metrics. *Am. Psychol.* **17**, 29–39.

STEVENS, S. S. & GALANTER, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *J. exp. Psychol.* **54**, 377–411.

STEVENS, S. S. & GUIRAO, M. (1963). Subjective scaling of length and area and the matching of length to loudness and brightness. *J. exp. Psychol.* **66**, 177–86.

TAYLOR, J. G. (1965). The behavioral basis of perceived size and distance. *Can. J. Psychol.* **19**, 1–14.

TEGHTSOONIAN, M. (1965). The judgment of size. *Am. J. Psychol.* **78**, 392–402.

(*Manuscript received* 5 *October* 1966)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 315–327
Printed in Great Britain

FURTHER STUDIES OF PATTERN PERCEPTION AND A STABILIZED RETINAL IMAGE: THE USE OF PROLONGED AFTER-IMAGES TO ACHIEVE PERFECT STABILIZATION

By C. R. EVANS

Autonomics Division, National Physical Laboratory, Teddington

A technique to allow the study of prolonged after-images as perfectly stabilized retinal images is described. A simple model to account for the variability of unitary or fragmentary disappearance of stabilized images is proposed and tested. Results suggest the presence of recognition units mediating the perception of straight lines which vary in size from approximately 30 to 90 min arc inthe human fovea and perifovea. These units appear to be arrayed independently of orientation, and may be primarily represented in the cortex. A reappraisal of the methodology and theoretical significance of stabilization studies is attempted.

When a pattern is stabilized on the retina, it may disappear in whole or in part (Ditchburn & Pritchard, 1956) and it has also been shown (Evans, 1965) that a significant percentage of fragmentation effects are non-random and that certain patterns may disappear more readily than others. Pioneer studies of stabilization, and practically all extended research until quite recently, has relied upon the contact-lens method of compensating for eye movements. This method depends upon attaching the target to be viewed to the contact lens itself, so that eye movements, voluntary and involuntary, are accompanied by simultaneous movements of the contact lens and the viewing apparatus. Thus, if contact-lens fit is good, the image of the target is held in one position on the retina. Studies of considerable interest to perceptual theory have been undertaken employing this method, but the system has at least two serious drawbacks which have made all researchers in the field eager to consider alternatives.

The first drawback concerns the difficulties experienced in inducing subjects to take part in experiments which require that they be fitted up individually with full-size contact lenses, and needs no further explanation. A more fundamental objection to the contact-lens method of stabilization is that put forward recently by Barlow (1963), who has pointed out that contact-lens slippage relative to the scleral surface may be appreciable (up to 30 min arc) and is certainly variable from trial to trial. Thus one is not always dealing with a perfectly stabilized image as has on occasions been tacitly assumed. Therefore, if one is to study very small or extremely simple targets, clearly either a running record of contact-lens slip must be made—a task which seems impossible—or some new method which achieves perfect stabilization must be found.

After considering Barlow's critique in which slippage of lenses had ingeniously been measured by comparing an after-image of the pattern against its 'stabilized' image, it was decided to examine the behaviour of some fairly complicated patterns viewed as after-images. The after-image, which is of course perceived as the result of signals generated by changes in the state of the retinal cells themselves, is to all intents

and purposes a perfectly stabilized image. Simple observational studies made it clear that patterns such as squares, triangles, circles with two cross-bars, and so on, could be seen clearly for an appreciable time when viewed as strong after-images, and could also be seen to fragment into the same units reported in contact-lens stabilization studies (Bennet-Clark & Evans, 1963). The perfectly stabilized after-image appears and reappears, sometimes in whole, sometimes in part, its length of stay being a function of the initial brightness of the flash, but the nature of its fragmentation being a function of the shape of the target—a finding in accord with earlier studies of stabilization phenomena, both qualitative and quantitative (Pritchard, Heron & Hebb, 1960; Evans, 1965). Despite this interesting finding it was decided not to pursue the study of relatively complex patterns, but rather, for the first extensive quantitative study using the after-image method of stabilization, to deal with much simpler units, in particular the straight line.

Previously (Evans, 1965) it had been shown that when a straight line is part of a more complex pattern there is a very high probability that this line will disappear or remain as a unit. Pilot studies of lines viewed as after-images seemed to suggest that fragmentary action increased as line length increased, and it was decided to investigate this further. In particular, the experiments were designed to examine the suggestions raised briefly in the earlier paper, that the fragmentation effects reflected the presence of classificatory perceptual units analogous to those discovered in the cat's visual system (Hubel & Wiesel, 1959). Now if in fact the tendency of a line to act as a unit does suggest the existence of perceptual units in the human system, then clearly the size and location of these units should have some bearing on the nature of the disappearance of stabilized images of straight lines.

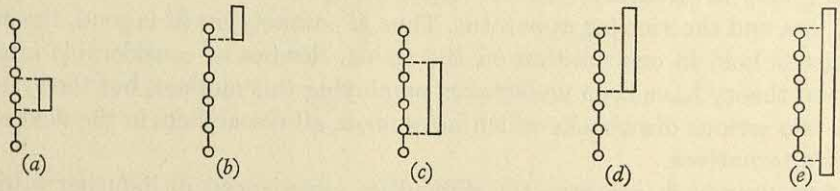


Fig. 1. A simple model predicting changes in unitary and fragmentary disappearance of an after-image as a function of the size of the after-image: (*a*) short line falling on unit produces unitary disappearance; (*b*) when line falls on edge of unit, fragmentation occurs; (*c*) and (*d*) as line length increases, probability of unitary disappearance falls; (*e*) when line is larger than unit, little or no unitary disappearance occurs.

The model

A simple model is now considered; it may be thought of as consisting of classificatory units which produce a rudimentary recognition (Fig. 1). The system is composed of a cluster of cells, and the primary rule is that the cells of any linked unit will have a greater probability of firing with each other than with any cell or cells outside the unit. Such a system would in fact constitute a very simple pattern recognition system for straight lines—such as might be built into an artificial retina by a very naive electronics engineer. It is now proposed that the image of a single vertical line, when falling on one of these receptive units, will be likely to be perceived as a unit, and when presented as a stabilized image will also tend to vanish as a unit. When

stabilized, however, for it to act in any way as a unit, it is reasonable to assume it to be necessary for the stimulus to fall within the local receptive field, since an image of given size overlapping one or more units not linked to synchronous firing will be more likely to fragment.

A simple series of predictions can be made from this model for the behaviour of stabilized images, particularly the perfectly stabilized after-images. The basic prediction is that if a given area of the retina is largely populated with receptive fields of a given size, and if stabilized images of varying size are produced, then as image size is increased, the probability of fragmentary disappearance will correspondingly increase. When image size is appreciably less than that of the local receptive fields, the probability of unitary action will be high. As image size approaches field size, the probability of fragmentary disappearance will steadily increase, and that of unitary disappearance steadily *decrease* until a value near zero is achieved.

Obviously this model does less than justice to the subtleties of retinal and cortical organization. Certainly fields subtending a degree or more in any part of the retina must involve very large populations of cells, and one would certainly expect such fields to be more than 'one receptor wide' as drawn in the diagram. Perhaps one might envisage a number of such units existing alongside one another, their ends overlapping to some extent, as constituting an effective receptive field. Provided that the basic units in a given area were of very similar size, the prediction that has been made for this introductory model would still hold.

Method

In principle the method is extremely simple; a pattern of the desired shape is illuminated for a brief duration by a powerful flash, and the subject then records his observations of the behaviour of the after-image by pressing one of a set of keys.

Flashing apparatus

For all the experiments described in this paper the brief illumination was provided by a Courtenay Major Studio-type electronic flash-head, capable of producing a variable flash output of up to 2500 joules. The flash source itself, a coiled gas-tube 10 cm in diameter, was encased in a polished aluminium tunnel, and a mirror was mounted to complete the back end of the tunnel. At the front end the tunnel was sealed off by a mask and a 2×2 in. glass slide holder for the targets. The patterns were prepared by carefully cutting the required shape out of the completely opaque surface provided by coating a 2×2 in. slide glass with black adhesive PVC tape. Thus any light emerging from the lamp passed through the target aperture, providing an intense and evenly distributed flash of 1 msec duration.

Procedure

Subjects were in all cases seated in a dimly lit room (room luminance was held constant at a level determined by measuring a constant patch with an SEI photometer) and were allowed to dark adapt for 15 min prior to each experimental run. At the standard level of luminance, the laboratory was sufficiently bright for newspaper print to be read without strain, though colours were markedly muted. Each subject sat in a wooden cubicle with black interior walls and rested his head on a firm support. A small aperture in the front of the cubicle allowed him to view the target stimulus.

In any given session, ten trials of 1 min each were given, 5 min being given between trials to allow the previous after-image to subside. The flash was viewed monocularly, alternate eyes being used from trial to trial.

After-images of straight lines or complex targets produced by this method are extremely clear,

and easy to observe and record with confidence. They undergo a cycle of appearance and disappearance, sometimes vanishing as a unit, and sometimes breaking into parts or more frequently losing one or other end. After-images in these experiments could be observed with some confidence for 2 min and with considerable clarity for the first minute, but always decayed before 5 min had passed.

Subjects pressed a key when unitary disappearance of the image was occurring, keeping the key pressed until the image reappeared in whole or in part. A second key was pressed when fragmentary disappearance of the image occurred. Subjects triggered the flash for themselves, after which they closed their eyes and kept them closed until the 1 min trial was ended. Recordings were automatically transferred to punched signals on the tape for computer analysis. After a pause following the conclusion of the trial, the subject would then fixate the stimulus with the other eye, press the key for the next flash and complete another run. Runs normally consisted of ten trials per person—that is to say, five after-images to each eye. Whilst there was some distinct trial-to-trial variation, the proportion of unitary to fragmentary disappearance remained reasonably constant for any one target size. There was no observable tendency, within the 1 min trials, for progressive changes to occur in the frequency of either category of disappearance (e.g. the probability of a fragmentary disappearance occurring was the same at the beginning as at the end of a 1 min run). All subjects who have taken part in repeated experiments have received frequent and regular ophthalmic examination, and as a precaution against possible retinal damage infra-red and ultraviolet glass filters were inserted between the flash tube and the target slide.

Results

Evidence relating to unit size in fovea

Fig. 2 illustrates the results of the first experiment (subject A; cf. Table 1, below) The flash was planted on the fovea. Ten 1 min trials were given for each of the five line lengths presented in all, and each symbol on the graph represents the mean of ten such trials (cf. Table 1, below). It is immediately clear that the results support the simple hypothesis advanced above. Fragmentary disappearance increases markedly and steadily as the retinal angle subtended by the after-image increases; conversely, unitary disappearance falls as line length rises.

Table 1. *Relations between percentage disappearance, either unitary or fragmentary, of after-images and length of target line (min arc) when after-image planted on fovea for four subjects (cf. Figs. 2 and 3 for subjects A and B)*

Subject	Disappearance (%)	Minutes of arc										
		10	20	30	40	(45)	50	(55)	60	70	80	90
A	Unitary	—	—	—	22·8	15·6	12·9	4·8	4·5	—	—	—
	Fragmentary	—	—	—	2·0	5·2	5·9	19·7	32·9	—	—	—
B	Unitary	—	17·8	4·9	0·3	—	2·3	—	0·7	4·0	—	—
	Fragmentary	—	22·3	35·7	31·0	—	34·3	—	38·0	39·3	—	—
C	Unitary	48·1	46·7	31·6	19·9	—	11·8	—	1·0	0·0	0·0	0·0
	Fragmentary	0·0	0·0	0·5	20·2	—	22·1	—	39·7	38·1	50·3	61·8
D	Unitary	20·2	24·6	15·4	8·5	—	13·1	—	10·9	12·6	1·3	0·0
	Fragmentary	0·0	5·8	14·7	32·8	—	21·1	—	29·9	36·3	40·4	36·9

It is by no means certain that the straight lines drawn in this and subsequent figures precisely express the slope and structure of the unitary and fragmentary curves; in several cases, however, it will be seen that a straight line does appear to be the most suitable, and for the sake of consistency the best fitting simple straight line has been calculated in each case. The cross-over points of the lines in this and

subsequent figures are of appreciable interest, for it can be shown, on the basis of our hypothesis, that the size of a receptive unit or field is approximately twice that indicated at the point of intersection of the lines (i.e. the point at which it is equiprobable that a fragmentary or a unitary disappearance will occur). Thus it appears that for subject A the size of these fields or units in the fovea is approximately 90 min arc.

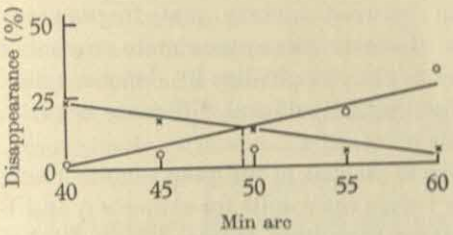


Fig. 2. Subject A: flash on fovea.

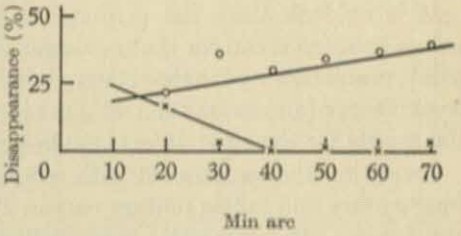


Fig. 3. Subject B: flash on fovea.

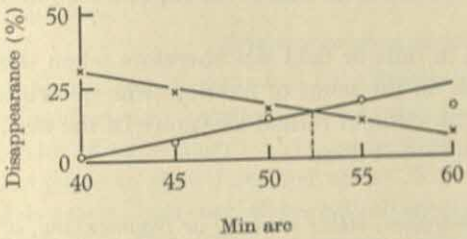


Fig. 4. Subject A: flash 1° off fovea.

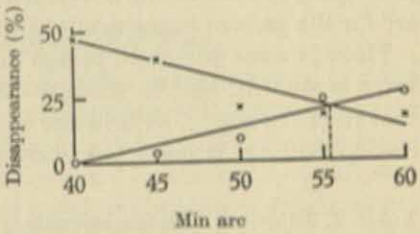


Fig. 5. Subject A: flash 2° off fovea.

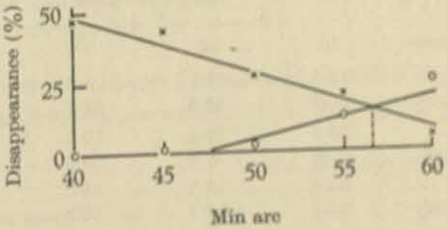


Fig. 6. Subject A: flash 3° off fovea.

Figs. 2–6. Relations between percentage disappearance of after-images, either unitary (×—×) or fragmentary (○—○), and length of target line (min arc) for different positions of the target line on the retina. Figs. 2 and 3 show that as the length of an after-image of a straight line is increased, percentage fragmentation increases and percentage unitary disappearance decreases; the cross-over points give information about the size of the hypothesized perceptual units. Figs. 2, 4, 5 and 6 show that planting after-images at increasing distances from the fovea systematically changes the relation between percentage unitary and fragmentary disappearance lines; the cross-over points also alter, suggesting a progressive increase in the 'field size' of units in the more peripheral parts of the retina.

The same experiment was then performed with subject B. Initially the same range of stimulus sizes were used (40–60 min arc), but preliminary analysis of the results indicated that while percentage fragmentary disappearance yielded the expected

rising curve, unitary values lay close to zero throughout the range. The inference that foveal field sizes in this subject were considerably smaller than those indicated for subject A could be tested by providing a further set of appropriately smaller after-images and the study was extended to include lines of 20 and 30 min arc in length. (In subsequent experiments with other subjects the greater range of image size has almost always been used.) Fig. 3 shows the results for this subject with foveal fixation (cf. Table 1).

It is evident that the predicted relation between unitary and fragmentary values holds as it did for the previous subject. However, the approximate cross-over point, even allowing for the relative imprecison of the straight-line fit, denotes a field or unit size of not more than 40 min arc, suggesting a significant difference between the results for the two subjects studied at this point.

Repeating the experiment with other subjects yielded in all cases similar rising fragmentary and falling unitary curves. Table 1 gives the results for subjects A and B and for two other subjects; very similar patterns have been obtained for a further six subjects. From person to person the estimated 'cross-over' point tends to vary somewhat, but indicating nevertheless that field sizes lie within the range 30–90 min arc for the present experimental situation.

There is some person-to-person variation in unit or field size therefore when the fovea is studied; that is, when the flash falls on the point of fixation. The question that next arises is: does field size vary as a function of retinal location? In the next section evidence is presented which suggests that it does.

Table 2. *Relations between percentage disappearance, either unitary or fragmentary, of after-images in different positions on the retina and length of vertical target line (subject B)*

		Length of target line (min arc)				
Displacement	Disappearance	40	45	50	55	60
1°	Unitary	16·3	11·5	8·6	6·1	6·8
	Fragmentary	40·9	44·9	51·7	57·9	66·4
2°	Unitary	33·9	26·4	19·1	13·2	11·7
	Fragmentary	37·6	46·9	53·4	63·1	65·7
3°	Unitary	36·4	26·7	23·0	14·5	15·5
	Fragmentary	43·9	52·2	59·3	67·3	68·1

Evidence relating to unit size in periphery

The experiments above involved foveal siting of the after-image. In the following experiments subjects fixated either 1°, 2° or 3° to the right or left of the flash, the fixation point being marked by a white dot. For subject A after-images were always positioned on the right hemiretina of either left or right eyes. For subject B a longer series was instigated in which both hemiretinae were plotted; that is, after-images were placed 1°, 2° and 3° to the right and left of both fovea. For each retinal site, as in the foveal experiments, ten 1 min trials were given for each line-length.

Figs. 4, 5 and 6 show the mean values for 1°, 2° and 3° displacements of the flash, respectively, for subject A. The pattern predicted from the simple model and found in all the previous studies is again to be seen quite clearly in the studies of

more peripheral regions of the retina. In general, the steepness of the ascending and descending curves varies with the change of retinal site of the after-image, and this may reasonably be interpreted as indicating a change in size of field or receptive unit. Allowing for the approximate nature of the straight lines expressing the data, field size increases with movement out from the fovea, the size varying within the probable range 100–120 min arc. Comparable data for subject B are given in Table 2.

With the greater number of trials involved in this phase of the experiment, the curves can be seen to have become smoother. The progressive rise and fall of fragmentary and unitary disappearance values respectively is maintained and in complete accord with prediction. It is evident that extrapolation of the curves should indicate an early cross-over, first of the 3°, then the 2° and finally the 1° lines. Approximate estimates for the position of the cross-over points are:

$$1° = 30 \text{ min arc}; \quad 2° = 35 \text{ min arc}; \quad 3° = 40 \text{ min arc}.$$

This tendency for indicated unit size to increase as the stimulus is moved into the periphery has been confirmed in observations on a further six subjects, and experiments currently in progress are dealing with stimuli displaced by 5° and 10° out from the fovea.

Evidence relating to orientation of receptive units

Earlier experiments had been designed to search for differential responses, in terms of unitary and fragmentary disappearance, to lines of varying orientation. A further series of experiments have been carried out using a method which was similar in principle to that described above. The experiments were carried out in a different laboratory, however, where overall luminance and flash brightness were unfortunately very dissimilar. Thus the precise values listed may not be compared directly with those detailed in the sections above dealing with field size. They may, however, safely be compared with each other, for conditions in the course of the orientation experiment itself were not altered.

Table 3. *Relations between percentage disappearance, either unitary or fragmentary, of after-images in different orientations on the retina and length of target line (subject B)*

		Length of target line (min arc)					
Orientation	Disappearance	40	45	50	55	60	90
12 o'clock	Unitary	0·0	0·0	0·0	0·0	0·0	0·0
	Fragmentary	48·9	54·3	55·6	60·8	63·9	66·6
1 o'clock	Unitary	0·3	0·0	0·0	0·0	0·0	0·0
	Fragmentary	47·7	52·1	58·3	56·3	70·1	57·9
2 o'clock	Unitary	0·9	0·0	0·0	0·0	0·0	0·0
	Fragmentary	57·9	55·6	57·1	59·1	62·9	69·6
3 o'clock	Unitary	0·4	0·0	0·0	0·0	0·0	0·0
	Fragmentary	52·0	57·2	58·4	58·0	64·6	65·5

Subject B was the principal subject in these experiments. He was presented with a series of flashes of a straight line of varying length in each of four orientations, termed '12', '1', '2' and '3' o'clock respectively. The subject underwent ten trials for each of six line lengths in each of the four orientations—a total of 240 1 min trials in all.

Table 3 gives the mean values for each of these conditions; the systematic rise of fragmentary disappearance observed in all previous studies is equally apparent in these results. More significantly for the present discussion, there was little, if any, change in the trend of the values whatever the orientation. This suggests strongly that for the simple pattern employed, field size does not change within this range of orientations.

Evidence suggesting post-retinal location of units

It has been remarked that the displacement of the after-image for subject B was systematically varied to allow, over a very large number of trials, equivalent exploration of the left and right hemiretinae of both eyes. This meant an extremely long experiment, but it was believed to be important because of the light a study of this kind might throw on the location, within the visual system, of the receptive units being studied.

The distribution and density of the receptors within the retina does not remain constant or symmetrical as one moves to the left or right of the fovea (Østerberg, 1935); rather, the nasal hemiretinae of the right and left eyes are more similar anatomically than are the nasal and temporal maps of any one eye. However, the right hemiretinae of both eyes in man feed into the right hemisphere and the left hemiretinae into the left hemisphere; hence anatomically similar retinal points feed to opposite hemispheres. It should be instructive therefore to re-plot the data obtained from subject B, pooled for both eyes in Table 2, according to the hemiretina involved. Data from additional studies involving a line subtending 90 min arc at the retina are also included in these graphs, each point on which represents the mean of ten 1 min trials.

Figs. 7 and 8 depict the data for unitary and fragmentary disappearance for 1° displacement on each hemiretina. It is immediately clear that the curves for each hemiretina are similar in shape to each other. Of greater interest, however, is the tendency, not completely clear-cut but evidently present, for the curves for the left hemiretinae (left hemisphere) to 'track' each other, and for those of the right hemiretinae (right hemisphere) to do the same. This effect appears to be more marked for the right hemiretinae curves. Figs. 9 and 10 present the data for the four retinal sites 2° off the fixation point. The 'tracking' noted above now emerges strongly, and once again this appears more clearly for the right hemiretinae. When the data for the 3° displacement sites are plotted (Figs. 11, 12) the effect is quite unequivocally demonstrated. Evidently unit size is unaltered when right hemiretinae are compared with each other, or, conversely, when left hemiretinae are compared with each other. This would suggest most strongly a post-retinal or cortical site for the units.

DISCUSSION AND CONCLUSIONS

Before drawing conclusions about the significance of these findings, it is necessary to consider the relations of the experiments to the entire previous series employing contact lenses which has been continuing in one form or another for a decade and a half. It would seem possible to argue that the results outlined cannot be taken to throw much light on the stabilization topic because the after-image is an essentially different type of visual phenomena. This objection needs to be met, and in order to

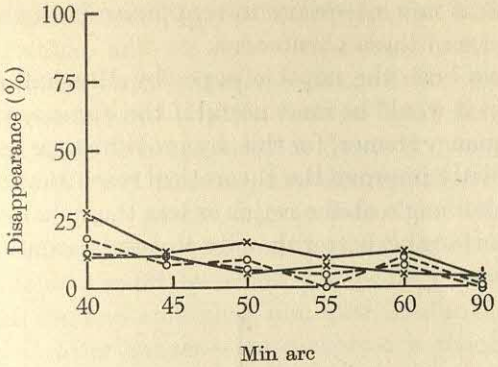


Fig. 7. 1° displacement from fovea: unitary disappearance.

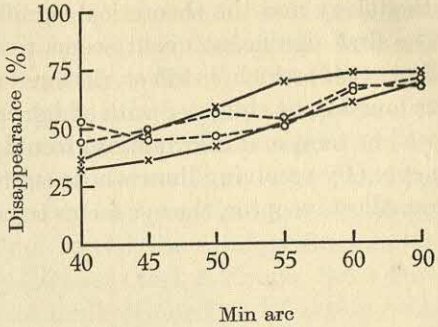


Fig. 8. 1° displacement from fovea: fragmentary disappearance.

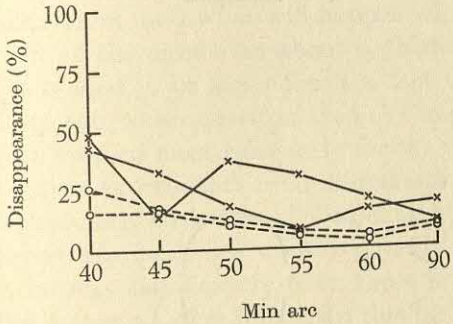


Fig. 9. 2° displacement from fovea: unitary displacement.

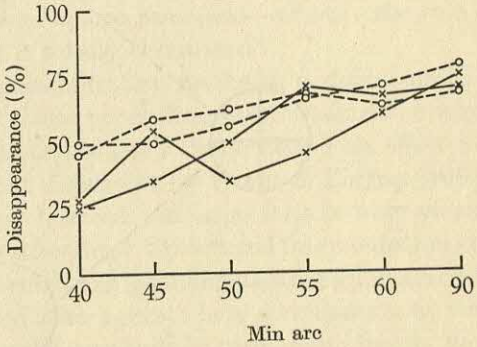


Fig. 10. 2° displacement from fovea: fragmentary disappearance.

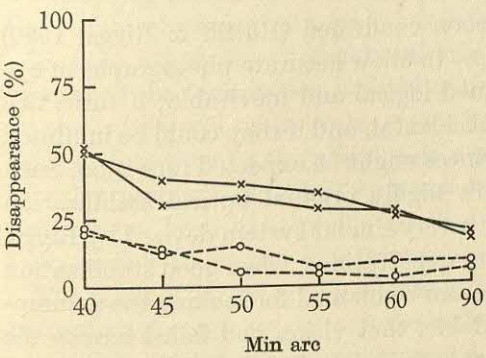


Fig. 11. 3° displacement from fovea: unitary disappearance.

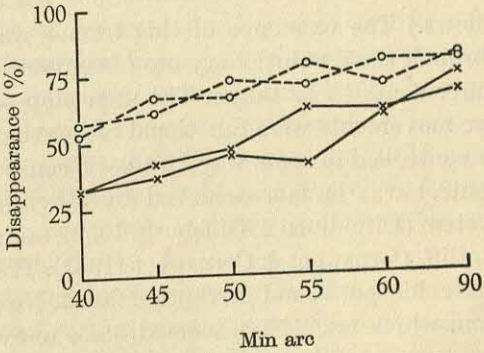


Fig. 12. 3° displacement from fovea: fragmentary disappearance.

Figs. 7–12. Relations between percentage unitary and fragmentary disappearances of after-images and length of a vertical target line (min arc) at different distances from the fovea in the hemiretinae of right and left eyes (subject B); ×—×, Values for left hemiretinae; O- - -O, values for right hemiretinae.

do this the history of investigations using stabilized images needs to be considered. It is suggested that the outcome is that it is now necessary to reappraise both the methodology and the theoretical significance of these phenomena.

The first significant event seems to have been the notable paper by Marshall & Talbot (1942) which, in effect, declared that it would be most useful if the human eye were found to be equipped with a high-frequency tremor, for this, by providing spatial as well as temporal information, would greatly improve the theoretical resolution of the eye. (By resolving lines whose subtended angle at the retina is less than that of the smallest receptor, the eye seems to be performing better than its physical potential

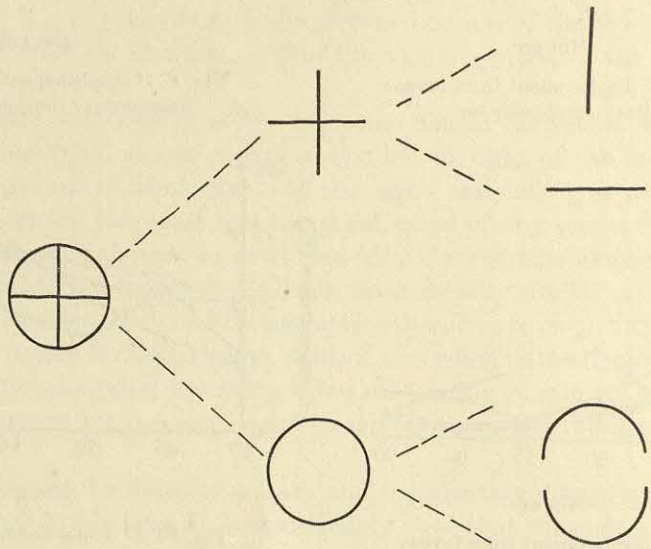


Fig. 13. A possible perceptual classificatory hierarchy.

allows.) The existence of this tremor was soon confirmed (Ratliff & Riggs, 1950) through the revolutionary use of contact lenses to allow accurate photographs of eye movements to be taken. The next step seemed logical and inevitable; if these fine eye movements were functional rather than accidental, and if they could be inhibited or controlled in some way, dramatic consequences might be expected for vision. Such control was in fact achieved by Ditchburn's highly original optical stabilization system (Ditchburn & Ginsborg, 1952), and by a very similar system devised by Riggs, Ratliff, Cornsweet & Cornsweet (1953). True to expectations, when good stabilization was achieved, visual perception failed. No one can be blamed for making the assumption, which must have seemed quite unavoidable, that vision had failed *because the fine tremor had been controlled*. This assumption, which is still widely held today, has led to some awkward complications and apparent contradictions that have tended to slow down the theoretical development of this subject.

The first attack on this position was made by Barlow (1963), who showed that comparison of an after-image with a contact-lens stabilized image revealed slippage of an extent which could hardly be ignored. Barlow, it is true, had only demonstrated slippage in his own lenses, and better stabilization might well be achieved in other

laboratories, but at least two new critical questions had been raised. First, was it justifiable any longer to experiment with contact lenses whose slip was, at best, variable, and to make deductions which assumed complete or at least constant stabilization? Secondly, was not the whole range of 'psychological' phenomena associated with stabilization suddenly far less convincing? Much of the interest surrounding these phenomena depends upon the partial nature of the disappearance of patterns and of the regeneration effects, and partial destabilization could quite easily cause much partial fragmentation and all forms of regeneration. A previous paper (Evans, 1965) attempted to counter this criticism by showing that poor stabilization might account for some, but could not account for all non-random fragmentation effects; and an earlier, non-quantitative study (Bennet-Clark & Evans, 1963) showed that after-images—presumably the theoretical limit of good stabilization—disappeared in non-random fashion. In apparent contradiction to this were the repeated observations of Yarbus (1957*a*, *b*), who used an uncomfortable but extremely adhesive lens and found that an image, once vanished, would not return unless it was re-illuminated in some way. The central issue of course remained—what is the role of fine tremor, and what will happen when it is totally eliminated?

In all the discussion about sophisticated stabilization methods, a significant fact has tended to be forgotten: the fact that visual perception may break down when there can be no question that the fine tremor has been eliminated. This effect has been studied most relevantly for the present discussion by Evans & Piggins (1963). In their experiments room illumination was lowered and large targets were viewed monocularly with the aid of a dental bite and headrest. Structured fragmentation and partial reappearance were frequently observed, the phenomena appearing to differ in no way significantly from those reported with contact-lens stabilization or with the patterned after-image. In this instance the involuntary tremor can hardly have been affected at all; with contact lenses it may have been partially compensated; while in the case of the after-image it must presumably have been totally eliminated. Whether any relation between disappearance or fragmentation effects and compensation for the physiological tremor has in fact been demonstrated becomes questionable.

To date, attempts to understand stabilization phenomena have clung to what might be termed the 'adaptation or habituation' approach. An image held in one position on the retina will cease to be seen after a period because local receptors have no chance to recover in the absence of a moving stimulus. The behaviour of the more complex patterns and their apparent natural break-up into 'meaningful' units has furthermore been unhelpfully ascribed as in some way supporting Gestalt theory. Is it possible to break away from these approaches?

The work of Hubel & Wiesel (1959, 1962), producing evidence for the existence of perceptual recognition units in the cat's visual cortex, has suggested that a search for something similar in the human visual system might be profitable. This has in fact been the rationale for the work described in this paper, and for other studies such as those of Andrews (1965). If, as seems reasonably likely, the human perceptual system works on a hierarchical classificatory basis, then one might hope under certain circumstances to 'catch the system at work'. A classificatory system such as that in Fig. 13 could easily be constructed to deal with the recognition problems of a very simple visual world. In any working system of this sort, whether electronic or bio-

logical, total perception (the activation of the topmost level of the hierarchy) will not be achieved without some transfer of information at each level in the system. Clearly, the more complex the system, the more information transfer is necessary, and, if some limits were to be placed on the information supply, the activation of only part of the hierarchy would be possible. An operation analogous to this in the human visual system emerges when one of the approaches to 'stabilization' is considered. In (*a*) steady fixation, or (*b*) partial stabilization by contact lens, or (*c*) total stabilization by the after-image, information about the pattern is restricted in such a way as to interfere with the classificatory process, and not all the 'switches' or levels in the hierarchy are activated. The individual would then see *parts* of the pattern, and the parts seen would correspond to the level of the hierarchy reached and to the nature of the classificatory units themselves. The level reached would fluctuate as information supply fluctuates, and occasionally (though momentarily only), the topmost level would be reached. The more complex the pattern, however, the more prone it would be to 'fragment' (cf. Evans, 1965), and the probability of fragmentation should be a function of the number of perceptual units involved in a total hierarchical system.

Why, then, should an after-image, once it has vanished, reappear at all if it is a perfectly stabilized image? The answer would seem to be that while the after-image is indeed spatially stabilized, information about its contours and form is being supplied, though in reduced amount, by the pigment regeneration processes occurring in the retina. When perfect stabilization is achieved, however, without the long-term processes involved in the after-image, one might indeed expect permanent fading without reappearances to occur. In this way the apparent discrepancies between the writer's findings and those of Yarbus could be comprehended.

In conclusion it can be predicted that if this hypothesis of what might loosely be called 'information starvation' is correct, it should be possible to engineer fragmentation effects by methods different from all those reported above. The writer believes that such a method has been found, involving, perhaps rather surprisingly, the tachistoscope as a research tool. When patterns of the type that have been studied as stabilized images are presented for brief periods in the tachistoscope and given repeated exposures without the exposure time being raised, subjects report the perception of the characteristic stabilization fragments. Work on this topic is at present in progress.

The work described has been carried out as part of the research programme of the National Physical Laboratory.

REFERENCES

ANDREWS, D. P. (1965). Perception of contours in the central fovea. *Nature, Lond.* **205**, 1218–20.

BARLOW, H. B. (1963). Slippage of contact lenses and other artefacts in relation to fading and regeneration of supposedly stable retinal images. *Q. J. exp. Psychol.* **15**, 36–51.

BENNET-CLARK, H. C. & EVANS, C. R. (1963). Fragmentation of patterned targets when viewed as prolonged after-images. *Nature, Lond.* **199**, 1215–16.

DITCHBURN, R. W. & GINSBORG, B. L. (1952). Vision with a stabilized retinal image. *Nature, Lond.* **170**, 36–7.

DITCHBURN, R. W. & PRITCHARD, R. M. (1956). Stabilized interference fringes on the retina. *Nature, Lond.* **177**, 434.

EVANS, C. R. (1965). Some studies of pattern perception using a stabilized retinal image. *Br. J. Psychol.* **56**, 121–33.

EVANS, C. R. & PIGGINS, D. J. (1963). A comparison of the behaviour of geometrical shapes when viewed under conditions of steady fixation, and with apparatus for producing a stabilized retinal image. *Br. J. physiol. Opt.* **20**, 1–13.

HUBEL, D. H. & WIESEL, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol., Lond.* **148**, 574–91.

HUBEL, D. H. & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol., Lond.* **160**, 106–54.

MARSHALL, W. H. & TALBOT, S. A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory activity. *Biol. Symp.* **7**, 117–64.

ØSTERBERG, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta ophthal.* (Suppl.), **6**, 76.

PRITCHARD, R. M., HERON, W. & HEBB, D. O. (1960). Visual perception approached by the method of stabilized images. *Can. J. Psychol.* **14**, 67–77.

RATLIFF, F. & RIGGS, L. A. (1950). Involuntary motions of the eye during monocular fixation. *J. exp. Psychol.* **40**, 68.

RIGGS, L. A., RATLIFF, F., CORNSWEET, J. C. & CORNSWEET, T. N. (1953). The disappearance of steadily fixated visual test objects. *J. opt. Soc. Am.* **43**, 495–501.

YARBUS, A. L. (1957*a*). A new method of studying the activity of various parts of the retina. *Biofizika* **2**, 165–7.

YARBUS, A. L. (1957*b*). The perception of an image fixed with respect to the retina. *Biofizika* **2**, 683–90.

(*Manuscript received* 3 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, 329–335
Printed in Great Britain

ADAPTATION TO PRISMS: CHANGE IN INTERNALLY REGISTERED EYE-POSITION

By B. CRASKE
Department of Psychology, University of Southampton

It was shown that subjects who inspected immobile parts of their own body through ophthalmic prisms subsequently pointed incorrectly towards visual targets with both hands, i.e. they pointed as if signals derived from the system responsible for positioning the eye had been changed by a constant. It was further shown by a direct method that the subjects' appreciation of eye-position had indeed undergone a change as a result of such inspection, and that it was of an order great enough to explain the observed changes in pointing behaviour.

Experiments on the adaptation of human and infra-human behaviour to the distortions caused by wearing laterally displacing prisms in front of the eyes go back 100 years, but experiments specifically designed to determine the mechanisms involved, as opposed to defining the content area for exploration, are of relatively recent origin. Held and his co-workers have undertaken a series of experiments exploiting the importance of re-afferent information as an underlying variable (Held & Freedman, 1963). A full summary of the field until 1965 can be found in Harris (1965). Recent interest has been directed towards an explanatory model involving a change in the 'felt position' of the limb used in the adaptation procedure (Harris, 1963). This may be conveniently thought of as a change in transfer function of output of joint receptors (Craske, 1966). However, it is important to note that this sort of explanation is basic only for the restricted situation where the observer views the target through prisms with the head immobile (Hamilton, 1964) and subsequently learns to point at it correctly, there being no delay between immediate knowledge of results, withdrawal of the arm and initiation of another attempt to point accurately. In other situations a change in the felt position of the head on the shoulders and of the eyes in the head have been put forward as possible contributing factors (Harris, 1965; Hamilton, 1964). Such a view seems reasonable in the light of recent examples of adaptation in a prism situation which are not explicable on the 'change in joint receptor transfer function' hypothesis (McLaughlin & Bower, 1965; McLaughlin & Rifkin, 1965; Hein, 1965). That changes in apparent orientation of the head on the body do occur is implied by Hein's 'postural after-effects', which show up as incorrect pointing to visual targets after the observer holds his head turned to one side for 10 min (Hein, 1965).

A change in appreciated eye-position has been inferred by Hay & Pick (1966) after the observers had viewed their moving body through prisms. Kohler (1964) directly observed changes in the eye-positioning system when, after wearing half-prisms (i.e. a conventional ophthalmic prism cut in half and worn over the lower half of the eye, base left or right), his observers produced a step-like eye-movement when asked to make one vertically, and were not aware of their mistake.

The experiments reported in this paper are a direct attack on appreciated eye-position in restricted and specifiable situations, and were carried out under condi-

tions which preclude adaptation due to change in transfer function of joint receptors and 'postural after-effects' mentioned above.

EXPERIMENT I

Method

Apparatus

Two twenty-dioptre prisms were mounted, base right, in a vertical board; attached to the bottom and at right angles to this board at mid-chest height was a horizontal board of sufficient length to allow the observer to see the finger of an outstretched arm. An inclined board running from the top of the vertical one to 2 in. above the far end of the horizontal one gave the observer a field of view which was restricted to a horizontal slot.

Table 1. *Summary of results of Expt I*

(Readings in degrees; scale origin the midpoint of eyes.)

	Mean error on pre-test		Mean error on post-test	
Subject	Right hand	Left hand	Right hand	Left hand
1	+3·2	+2·9	+6·7	+7·2
2	+0·6	+0·9	+6·1	+3·9
3	−2·1	−1·5	+3·3	+4·1
4	+1·0	+1·7	+8·6	+8·4
5	+1·2	+0·3	+2·4	+6·3
6	−1·9	−1·6	+9·1	+8·9
7	−2·4	−2·0	+5·4	+5·0
8	+2·6	+2·1	+8·2	+6·6
9	+1·3	+1·0	+5·7	+6·0
10	+1·1	+1·9	+9·8	+7·2
Mean	+0·56	+0·57	+6·53	+6·36
s	1·836	1·636	2·325	1·588

The 95 % confidence intervals for the above means are respectively: $0{\cdot}56 \pm 3{\cdot}58$; $0{\cdot}57 \pm 3{\cdot}21$; $6{\cdot}53 \pm 4{\cdot}56$; $6{\cdot}36 \pm 3{\cdot}11$.

Pooling the results for hands before and after exposure, $t = 8{\cdot}33$, D.F. 9, $P < 0{\cdot}0005$.

In another experiment, using different subjects, the same significant shift was shown when the observer inspected the index finger of the left hand during the treatment condition.

Procedure

The prisms were adjusted to the interpupillary distance of the observer and then swung aside. The observer was adjusted for height using a hydraulic chair, and was asked to bite on a dental-wax mouthpiece. Any tendancy to body-swivel was controlled with a trunk holder. These devices did not interfere with reaching movements of the arms.

Pre-test

A single-line target was moved into one of five positions symmetrical about the visual axis and in such a way that the observer could see it through the slot. The observer pointed at the target with right and left hands alternately for a total of twenty pointings with each each hand. Throughout, his arms were under the horizontal board and thus he was without knowledge of results. Viewing took place without prisms.

Treatment

The prisms were swung into position and the observer extended his right arm beneath the horizontal board in such a way that the tip of his index finger could be seen to be in the middle of the prism-altered visual field. The whole arm was supported on a moulded polyether pad, and strict instruction was given to hold the limb quite immobile during the ensuing inspection period. The observer inspected his immobile fingertip for 3 min.

Post-test

After this inspection the prisms were swung away, and five pointings with both right and left arm were taken, as in the pre-test. The treatment condition was then repeated, followed by the post-test until a total of twenty post-inspection pointings had been taken for each hand. This procedure was adopted in order to guard against the rapid decrement of prism-induced effects with time (Hamilton & Bossom, 1964). All error readings were in degrees, the circular scale, origin between the eyes, being projected on to the straight edge of the horizontal board.

Results

The data show significant shifts in errors of localization after the treatment for both the right and left hand, but more importantly the inspected hand showed no greater shift than the uninspected. This makes plausible the idea that the position information signalled by the eye might have changed. As a direct attack on this problem, the following experiments were carried out.

Experiment II

Preliminary experiments had indicated that the observer inspecting his own feet through prisms was perhaps more effective than inspecting the finger in generating the pattern of results observed in Expt I, and this method was used in Expt II. This procedure had the effect of increasing the time delay between treatment and post-test; thus, in order to offset any decrement the effect might show with time (Hamilton & Bossom, 1964), exposure time was increased to 10 min in the hope either of raising the level of prism adaptation or of making the adaptation more resistant to the decay process.

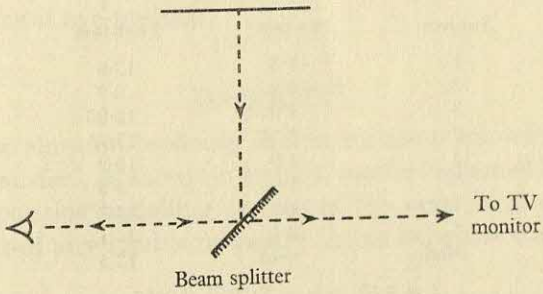


Fig. 1

Method

Apparatus

The observer sat with his right eye in front of a TV camera, the lens of which was behind an 8 × 6 in. semi-silvered beam-splitter which was inclined at an angle of 45° to the horizontal (Fig. 1). Reflector spot-lamps uniformly illuminated a smooth, featureless ceiling patch to give a brightness of 250 ft.-lamberts. This served the dual role of illuminating the observer's eye with sufficient light to get a good image on the TV monitor, and at the same time allowing the observer to see only the ceiling patch. When in position on the bite, the brightly illuminated ceiling was all the observer could see over the limits of an 80° horizontal arc. Behind the mirror, but not obscuring the eye from the camera, was a 100 W quartz–iodine lamp, some 9 in. from the observer's right eye. The TV hook-up gave a magnification of × 10.

Procedure

Pre-test. The observer sat in position with his head on the bite and was asked to look straight ahead. When he had achieved this to his own satisfaction, the observer indicated to the experimenter, who then drew a fine line tangentially to the left-hand side of the image of the observer's iris directly on to the TV monitor. The observer was then asked to look away and the procedure was repeated. Ten readings were taken.

Treatment. In this condition the observer wore trial frames with 20-dioptre base left prisms centred in front of each eye. He stood with his head symmetrical about the median sagittal plane and inclined such that his toes could be seen; the observer's head was fixed, using a bite bar and dental wax. The observer was asked to clasp his hands behind his back in order that they would not be seen. A small light was so placed that only his feet were illuminated, and all other room lights were extinguished. The observer inspected his toes for 10 min; no movement was allowed.

Post-test. Immediately following the treatment condition, the observer repositioned himself in the bite in front of the TV camera. Before looking straight ahead, however, the observer was asked to look at the quartz–iodine lamp, which was switched on for 10 sec. This was in order that any visual cues that the observer may have found in the pre-test would be masked by the light adaptation to the intense source. This procedure produces a strong after-image which was not present in the pre-test. It was found, however, that such an after-image did not impair and, as far as could be determined, did not affect the observer's eye-centring. Thus measures of eye-centring for an observer who has not been exposed to lateral visual displacement are drawn from the same population, regardless of whether an after-image is present or absent. Thus the purpose of the adaptation light is that of a precautionary measure; namely, masking possible cues. With a field of view which was perfect (a large area of white plastic, for instance) this precaution could be dispensed with. After looking at the light source for 10 sec, it was switched off and ten more measures of the voluntary straight ahead position were taken.

Table 2. *Mean difference in eye-position (in mm.) taken up before and after treatment as measured on the TV monitor*

(Zero is arbitrarily the first reading of each pre-test)

Subject	Pre-test	Post-test
1	−1·5	12·6
2	−0·9	9·2
3	1·9	15·95
4	5·75	17·75
5	1·0	13·2
6	−1·9	9·5
7	1·4	14·7
8	2·1	15·1
Mean	0·46	13·5

$t = 9{\cdot}57$; D.F. $= 7$; $P < 0{\cdot}0005$.

Results

All these shifts were in the direction of the adaptive shift shown in Expt I, which used pointing as the measure.

If the assumption is made that the limb reflects the positional errors derived from the eye on a one-to-one basis, the possible effect on pointing which a change in appreciated eye-position of this magnitude can produce may be calculated. To make this estimate, certain other data were required.

An observer was asked to scan along a metre scale at a distance of 65 cm (arm's length). A scan of 16 cm was shown to generate a shift of 3 cm on the TV monitor. Using the mean difference of the readings above as an estimate of change in the

position signal generated by the eye, it would be expected that 10 min eye adaptation would give a mean effect on pointing of $(1{\cdot}3 \times 16)/3 = 6{\cdot}9$ cm. It is of interest to note that this is not incompatible with the results of Expt I. (To determine the mean result of Expt I in centimetres, the mean difference of 5·88° in pointing and an estimated arm length of 65 cm yield a derived equivalent displacement of 6·5 cm.)

It might be argued that the above results were not due to the prism-exposure situation at all, but due to the fact that, in looking at his toes for 10 min, the observer had to hold his eyes at an angle of 11° to the straight ahead, and that the observed results may be due to this prolonged asymmetry of eye-position.

Experiment III

Expt III tested this proposal by two methods.

Procedure

(*a*) The procedure was the same as that employed in Expt II except that in the treatment condition a mark on the floor just forward of the feet was used as the target for inspection rather than the feet themselves. This spot was illuminated, no part of the observer's body was visible and the observer fixated the spot for 10 min.

(*b*) Again as for Expt II, but the treatment consisted in the observer remaining in the bite in front of the TV camera, and turning his eyes to maximum comfortable asymmetrical convergence, and there fixating a rod for 10 min.

Results

Using six observers, both these attempts failed to give any differences between pre- and post-treatment measures of the position of the observer's eyes when putting them straight ahead. The results were as if drawn from the same population.

It is concluded, therefore, that the change in internally registered eye-position is not due to asymmetrical eye-position.

Discussion

The results seem to show quite clearly that in certain prism-adaptation situations the 'inter-manual transfer', as shown in Expt I, can be explained by a change which takes place in the position-signalling system of the eyes. It is also clear that no voluntary movement of any kind is necessary in the exposure condition to produce this change.

The relationship between the positional errors signalled by the visual system and errors of localization in reaching and pointing is not known; the assumption that it is in a ratio of one to one gives comparable results for Expts I and II. In that most subjects have a reasonably large error range when pointing without visual guidance of the limb, a large number of readings would be necessary to make definite statements about this relationship.

Since these experiments were carried out, Kalil & Freedman (1966*a*, *b*) have presented evidence that there is eye-involvement in what has become a conventional prism-adaptation situation, namely the observer pointing to a visual target while wearing prisms, and only getting knowledge of results when the movement is complete. Previously, it had been accepted that there was no significant 'inter-manual transfer' of adaptation (Harris, 1963; Mikaelian, 1963; Hamilton, 1964) and hence,

by inference, no involvement of the eye-positioning system. However, ongoing experiments by Mr A. Johnston in the psychological laboratory at Durham have confirmed that a change in pointing of the unused arm does occur in this conventional situation, and that this is related to the magnitude of the change in appreciated eye-position as directly observed with the TV hook-up used for Expts II and III. This finding is in line with that of Rock, Goldberg & Mack (1966), who have shown significant correction to prismatic distortion when the exposure condition is limited to mere inspection of a prismatically distorted scene. This eye-involvement in a situation which primarily changes the felt position of the limb (Harris, 1963; Craske, 1966) clearly cannot be very great. Under these circumstances adaptation in the eye-positioning system can be of an order no greater than is necessary to 'take up the slack', i.e. can be no greater than that necessary to adapt to any residual error after the very rapid adaptation at the joint.

Evidence is being gathered in this laboratory that the population of observers is not uniform with respect to the readiness with which appreciation of the eye-position will change. This is only to be expected, but so far no account has been taken of this in the explanation of adaptation to prisms. In view of this, it would be expected that degrees of eye-involvement after exposure to a prism situation would depend upon the readiness with which both joint and eye will modify position data, which of these systems is being favoured by the exposure situation, and the degree to which the system and subsystems can tolerate discrepant positional information. Such a proposal would explain why Hay & Pick (1966) found no changes in appreciated eye-position subsequent to a prism-treatment condition where the observer viewed his hand sorting playing-cards. That is, this situation may favour adaptation in the joint and leave residual error too small to initiate a change in the eye system. Certainly their conclusion that there was not enough intersensory conflict to get visual adaptation was not justified, as Expt I illustrates. Also, since these experiments were carried out, Cohen has indicated the possibility of yet another 'eye factor' in the explanation of adaptation to prisms, namely a change in retinal space values (Cohen, 1966). His results have been confirmed by B. P. Moulden (personal communication). The next step in this area must be to define the necessary and sufficient conditions to generate these different factors involved in explaining adaptation to prisms, namely change in felt position of used limb, change in appreciated position of eyes and, if established, change in retinal space values.

CONCLUSION

Viewed together, these experiments constitute a case for the occurrence of a change in internally registered data about eye-position. The nature of the change will require an elucidation of the eye-position feedback system. A technique based on the work reported here has given results which might help to develop ideas about the appropriate mechanism (Craske & Templeton, in press). Until this work has progressed further, however, mere speculation on the nature of the change would not be fruitful.

References

Craske, B. (1966). Change in transfer function of joint receptor output. *Nature, Lond.* **210**, 764–5.

Craske, B. & Templeton, W. B. Prolonged oscillation of the eyes induced by conflicting position input. *J. exp. Psychol.* (in press).

Cohen, H. B. (1966). Some critical factors in prism adaptation. *Am. J. Psychol.* **79**, 285–90.

Hamilton, C. R. (1964). Intermanual transfer of adaptation to prisms. *Am. J. Psychol.* **77**, 457–62.

Hamilton, C. R. & Bossom, J. (1964). Decay of prism after effects. *J. exp. Psychol.* **67**, 148–50.

Harris, C. S. (1963). Adaptation to displaced vision: visual, motor, or proprioceptive change? *Science, N.Y.* **140**, 812–13.

Harris, C. S. (1965). Perceptual adaptation to inverted, reversed and displaced vision. *Psychol. Rev.* **72**, 419–44.

Hay, J. C. & Pick, H. L. (1966). Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *J. exp. Psychol.* **71**, 150–8.

Hein, A. (1965). Postural after-effects and visual-motor adaptation to prisms. (Paper read at E.P.A., Atlantic City.)

Held, R. & Freedman, S. J. (1963). Plasticity in human sensorimotor control. *Science, N.Y.* **142**, 455–62.

Kalil, R. E. & Freedman, S. J. (1966*a*). Intermanual transfer of compensation for displaced vision. *Percept. mot. Skills* **22**, 123–6.

Kalil, R. E. & Freedman, S. J. (1966*b*). Persistence of ocular rotation following compensation for displaced vision. *Percept. mot. Skills* **22**, 135–9.

Kohler, I. (1964). The formation and transformation of the perceptual world. *Psychol. Issues* **3** (Monogr. 12).

McLaughlin, S. C. & Bower, J. L. (1965). Auditory localization and judgements of straight ahead during adaptation to prism. *Psychonom. Sci.* **2**, 283–4.

McLaughlin, S. C. & Rifkin, K. I. (1965). Change in straight ahead during adaptation to prism. *Psychonom. Sci.* **2**, 107–8.

Mikaelian, H. (1963). Failure of bilateral transfer in modified eye–hand co-ordination. (Paper read at E.P.A., New York.)

Rock, I., Goldberg, J. & Mack, A. (1966). Immediate correction and adaptation based on viewing a prismatically displaced scene. *Percept. Psychophys.* **1**, 351–4.

(*Manuscript received* 16 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 337–350
Printed in Great Britain

STEREOSCOPIC VISION AND ANISEIKONIC LENSES. I

By G. N. SEAGRIM
Australian National University, Canberra

A series of experiments was conducted in which aniseikonic lenses were used. The course of events occurring when a lens is worn was recorded and measured by means of a tracking task. The well-known delay in effect was found but the occurrence of perceived distortions opposite in direction to those predicted was also revealed. The findings are discussed in relation to existing ideas concerning depth perception and a suggestion is made for additions to these ideas by means of which the obtained results might be understood.

Aniseikonic lenses, first described by Ames, Gliddon & Ogle (1932), possess certain important advantages over other apparatus (haploscopes, stereoscopes, etc.) for the examination of binocular space perception. If properly manufactured (Ogle, 1950) their effects on perceived spatial locations can be predicted accurately on the basis of existing theories of stereoscopic vision (Ogle, 1950, 1962) and, in addition, they can be worn by anyone, at any time, and for as long as required without undue discomfort or inconvenience. It therefore becomes possible to conduct experiments under conditions which otherwise approximate to normal. Originally designed for the correction of severe cases of clinical aniseikonia (Lancaster, 1938), their experimental possibilities were quickly appreciated by Ames and his colleagues, but the number of studies carried out with them by psychologists has been small.

This lack of interest may stem from the use first made of them by Ames and others. Without wishing to detract from the significance of these studies (Ames, 1946; Gardner *et al.*, 1959; Martin, 1954), whose general validity and psychological relevance may be beyond question, it is clear that not enough is known about the general consequences of the use of the lenses to warrant some of the theoretical implications drawn from the results obtained. The extensive studies conducted by Ogle (1950, 1962) appear to have been of greater interest to optometrists than to psychologists.

The advantages offered by the lenses for more psychological studies are offset by the disadvantage that their use may involve the intervention of associated variables whose nature is not fully understood and which cannot, therefore, be taken into consideration in the interpretation of results. Ogle discusses these only briefly, while Ittelson (1960) recognizes the advantage of their presence for the understanding of functional aspects of binocular vision. The present series of investigations attempts to identify some of these variables.

Aniseikonic lenses

The properties of, and prescriptions for, aniseikonic lenses are fully described by Ogle (1950). For present purposes they may be described as follows. Consisting essentially of sections of cylinders ground so that the surfaces are parallel but of unequal curvature, they magnify in one meridian only. They possess some power but to an extent which Ogle considers to be unimportant and which he specifies (1950, p. 124). When a lens is worn in the 90° position (with the axis of curvature vertical) the image on the eye covered by the lens will be magnified in the horizontal meridian only and to a degree depending on the curvature of the lens and on the thickness and the refractive index of the material of which the lens is constructed.

A lens placed in the 90° position in front of one eye has the effect of altering all the existing horizontal relations between the two retinal images. This is represented schematically in Fig. 1 (*a*) and (*b*), in which F is fixated with symmetrical convergence. The retinal image of F is not shown as displaced, the visual direction of F being assumed to be normal to the surface of the lens. The visual direction of P is displaced to that of P' and, if the values of the parameters involved are known, the extent and direction of the apparent displacement of P (to P'') relative to F can be predicted on the basis of existing theories of stereoscopic vision. These predictions can then be compared with obtained effects, as has been done under a wide variety of conditions by Ogle (1950, 1962).

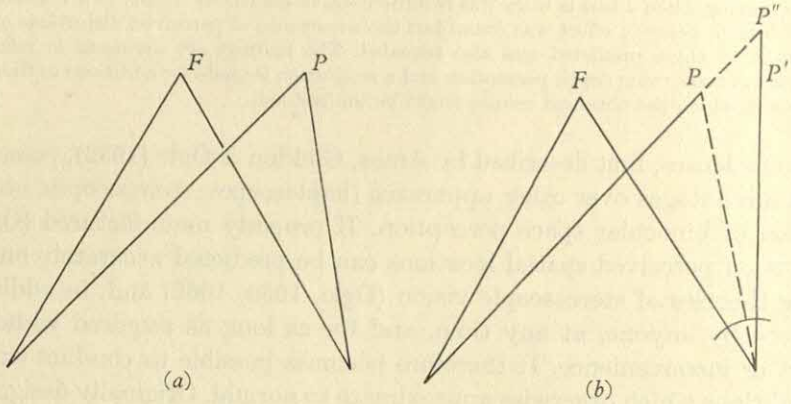


Fig. 1 (*a*) and (*b*). See text.

Three matters of obvious psychological interest have been reported by all authors working with these lenses. The first of these is the delay, or latency, with which the effects of the lens become apparent to the wearer. The second is the degree to which different individuals are affected by a lens. The third is the relation between the extent of the effect of a given lens and the situation in which it is worn. None of these matters appears to have been investigated adequately (but see Gillam, 1964). The first of the three, the latency with which the effects of the lens become apparent to the wearer, and individual differences in it, has sometimes been used as a dependent variable in personality studies (e.g. Gardner *et al.*, 1959). The course of events involved and the categories of possible response have not, however, been described in any detail. Ittelson (1960) reports some investigations but does not describe his findings in great detail, nor does he give an adequate description of his method of measurement (cf. Ittelson, 1960, p. 145).

EXPERIMENT I

This experiment was conducted to obtain measurements of events which occur when an aniseikonic lens is worn over one eye.

Apparatus

A tracking task was devised, the apparatus for which is shown schematically in Fig. 2. Two trolleys, to which a variety of objects could be attached, were mounted on guide rails, placed on the floor. The rails were aligned so that they converged on a point perpendicular to the midpoint between the subject's entrance pupils and subtended a constant visual angle of 7° 50′: physically, they diverged as they receded. The trolleys were controlled by cords which passed round pulleys

to wheels mounted on a table. These wheels carried helical grooves and enabled subject and experimenter respectively to alter the positions of one of the trolleys each. In this case the subject controlled the right-hand trolley.

The subject sat on a stool of adjustable height with his chin supported in a cupped rest. In this position he could turn the handle controlling the movement of the trolley without moving his head. The chin-rest was adjustable vertically and sagittally. It was mounted on a framework of metal rods on which were mounted two lens-holders. These were independently controlled by the experimenter with foot-pedals by which they could be placed before the subject's eyes smoothly and quickly.

A constant-speed paper and pen-recorder recorded the positions of the two trolleys independently. Event markers recorded the introduction or removal of the lenses. The way in which the subject moved his target to keep it apparently aligned with the experimenter's was thus recorded as a function both of time and of the presence or absence of a lens.

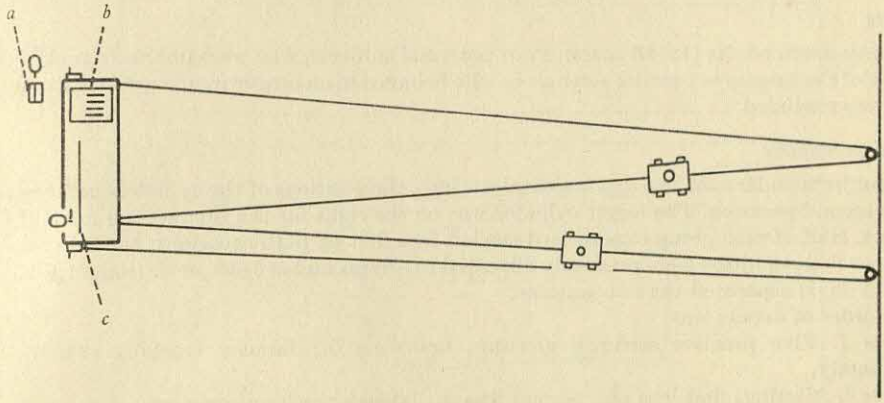


Fig. 2. Apparatus: a = experimenter's foot-pedal controls; b = pen-recorders; c = reduction screen in front of the subject.

Screens were arranged so that the subject could not see the trolleys but could see the upper parts of the target objects mounted on them, to a declination of approximately 10° 18′ to the horizontal plane. The targets, black metal cylinders 1·9 cm and 7·2 cm in diameter and 2 m and 1·8 m in height respectively, could be displaced from 2·60 m to 4 m from the subject. At 4 m distance they were 45 cm from the back wall. No attempt was made to screen parts of the room from the subject's view. Lighting was by daylight from the right, through windows and, only very occasionally, by overhead artificial incandescent lights. The total visual field can thus be described as normal.

Three percent lenses* were mounted on stalks and visually aligned to the 90° position. The height of the subject's head was adjusted to the lenses so that he reported viewing the whole of the visible parts of the targets through the lens when the latter was in position. This was done with monocular viewing. When withdrawn, the lens was completely clear of his effective visual field.

General experimental technique

Certain general procedures will be described and some terms defined.

Familiarization. The subject was familiarized with the apparatus and with the effects of the lenses. This was done by giving him a lens and asking him to examine the edge of a table-top through it until he experienced the induced slant. As latency was being examined, it seemed better for the subject to be aware of the nature of the expected response and it was hoped thus to reduce or to eliminate the contribution to latency of a simple delay in responding to what had

* Manufactured by the Optical Prescriptions Spectacle Makers, Sydney, Australia, to Ogle's (1950) prescription, under the supervision of Mr W. Bushby.

been noticed. Familiarization thus served as instruction, and the subject was also instructed how to adjust his trolley's position to compensate for the effects of the lens while the experimenter's trolley remained stationary.

Practice settings. Binocular settings were made at various distances, without lenses. The mean of such settings established the base-line for each subject against which experimental settings were measured.

Alerting procedure. While the targets were perceptually equidistant from the subject, the experimenter displaced his trolley about 30 cm and returned it to its position immediately before introducing the lens. The subject, who was informed about the procedure, did not track this movement which was supposed to attract his binocular attention.

Lens in and lens out. Movements of the lenses as they were placed in front of, or withdrawn from, the subject's eye.

Latency tracking. Adjustments made by the subject to the position of his trolley in response to the apparent displacements induced by the lens. Fixation was not demanded.

Subjects

Twenty-seven adults (18–35 years). Four potential subjects, who were unable to make reliable settings of the targets in practice settings or who behaved in an erratic manner during familiarization, were excluded.

Special procedures

All subjects underwent two experimental sessions, the positions of the cylinders being reversed on the second occasion. The larger cylinder was on the right for the first session for half of the subjects. Half of each group experienced the left lens first on both occasions; the other half the right lens first. Subjects were randomly allocated to groups and at least 24 hr (range: 1–25 days; mode: 5 days) separated the two sessions.

The order of events was:

Phase 1. Five practice settings; alerting; first lens in; latency tracking (2 min). Then immediately,

Phase 2. Alerting; first lens out, second lens in; latency tracking; lens out.

Results

Classes of response

Two distinct classes of response occurred. The first, primary latency, is that reported by all previous authors: a variable period of latency, during which the subject made no adjustment of the apparatus, followed either by an adjustment to the full displacement achieved, or by a gradual achievement of that by successive approximations. The second, labelled reversals, was a hitherto unreported response in which the initial adjustment was made in the direction opposite to that predicted.

Two classes of such reversed adjustments occurred. The first, when the subject turned the wheel in the wrong direction, is of no interest. This class of event was easily identified on the records by the rapidity with which it was corrected. Where any residual doubt existed, the subject was asked whether the movement had been in error or in response to a perceived change, and no subject had any difficulty in

---

Legend for Fig. 3.

Fig. 3. Tracings of records obtained in Expt I showing (*a*) short latency; (*b*) longer latency; (*c*) reversal.

Key. The vertical axis represents extent of displacement, the horizontal axis represents time. *S*, track of the subject's target displacements; *E*, track of the experimenter's target displacements; 1, practice settings; 2, alerting procedure; 3, right lens in; 4, left lens in; 5, end of primary latency, first lens; 6, end of secondary latency; 7, end of primary latency, second lens.

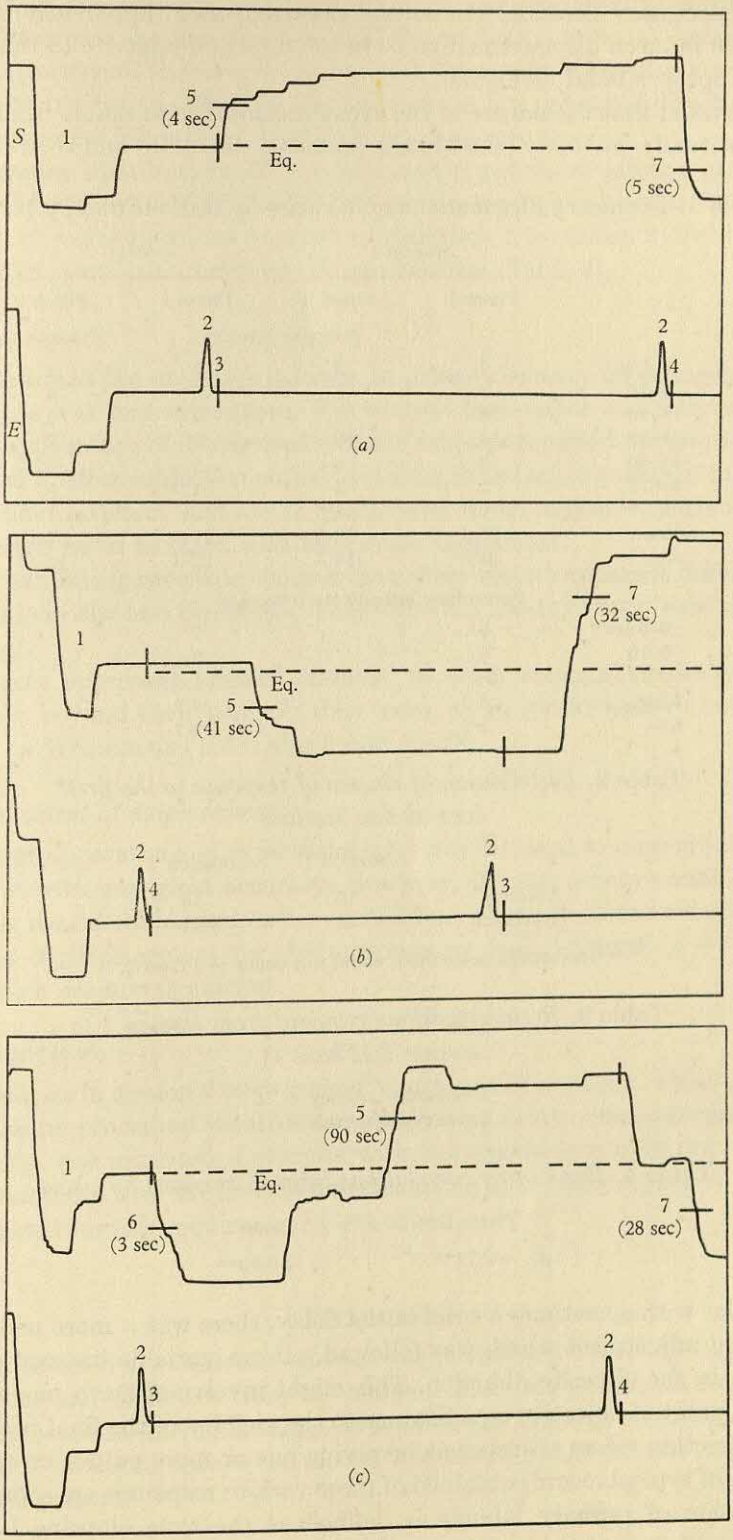


Fig. 3. For legend see opposite page.

making the necessary decision. The second occurred when, apparently, the subject perceived the induced displacement to be in the direction opposite to that predicted and made appropriate adjustments.

Subjects stated that the nature of the event was indistinguishable from the one to which it eventually gave way (slant in the predicted direction) and their responses to

Table 1. *Frequency distributions of latencies by sessions and by phases*

	Session 1		Session 2	
	Phase 1	Phase 2	Phase 1	Phase 2
	Primary latency			
0–5 sec	7	2	3	3
6–10	5	5	3	4
11–20	2	2	6	4
21–30	5	10	3	7
31–40	3	4	2	5
41–70	2	3	1	3
> 70	3	1	4	1
Mean	26·2	26·5	27·6	27·4
S.D.	28·8	18·5	32·3	23·7
	Secondary latency (to reversals)			
0–5 sec	6		7	
6–10	6		3	
10–13	—		1	
Mean	6·3		5·9	
S.D.	2·8		2·7	

Table 2. *Distribution of classes of response to the first* lens on two sessions*

	Session 1	Session 2
Latency	15	16
Reversal	12	11

* Reversals, as defined, could not occur in Phase 2.

Table 3. *Reclassifications required from session* 1 *to session* 2 (*Phase* 1)

Reversals–Latency	5
Latency–Reversals	1

Table 4. *Rank-order correlations between sessions by phases*

	Phase 1	Phase 2
ρ	+0·74**	+0·89**

it were similar with sometimes a brief initial delay, there was a more or less gradual and extended adjustment which was followed, after a variable interval of time, by a correction in the opposite direction. This might involve either a movement right through the point of objective equidistance to the vicinity of the final position in the predicted direction, or an adjustment involving one or more pauses on the way.

Examples of typical records obtained of these various responses are shown in Fig. 3. The duration of primary latency is defined as the time elapsing between the

introduction of the lens and the attainment of half the maximum displacement; secondary latency as the time to attain half the maximum reverse displacement. The criterion for classification as a reversal was that the extent of the displacement should exceed one-quarter of the ultimate displacement in the predicted direction and that it should persist for more than 5 sec (from start to recovery).

The frequency distributions of latencies and of classes of latency are set out in Tables 1 and 2 respectively. The reliability of response to the lens can be assessed by the number of reclassifications required from session 1 to session 2 (Table 3) and by the rank-order correlation between the two sessions (Table 4).

Introspective reports

These confirmed the objective records. In primary latency no change was noticed when the lens was first introduced. The targets then either suddenly or gradually took on an appearance of disalignment. Some subjects reported apparent movement, others denied it. Most subjects reported that their initial adjustments of the apparatus 'seemed to have no effect' and had to be repeated. Some subjects reported oscillations which were too rapid to track with the present apparatus.

Subjects exhibiting reversals claimed that there was no apparent difference in the displacements in the two directions. They were not aware of any 'cue' which might be operating.

All subjects expressed amused interest in their changing experiences, which were entirely beyond their control: they were, so to speak, passive but interested spectators, as Wheatsone (1852) also found (p. 23).

Latency and extent of displacement

The present apparatus and experiment were not designed to measure the extent of displacement with any great accuracy. However, as inter-person variability in both respects was considerable, rank-order correlations were calculated between latencies and extents of displacement for the first lens on both sessions: $\rho = +0{\cdot}217$ and $+0{\cdot}371$, which are not significant.

Effects of experience

Mean latencies in session 2 were greater than those in session 1. This indicates that the familiarizing procedure succeeded in eliminating delays due to surprise or to not knowing what was expected. Latencies with the second lens were not significantly different from those with the first. Before discussing the results in general any further, this unexpected finding requires to be examined.

Experiment II

It will be recalled that the second lens was introduced when the targets were at physically unequal distances from the subject. The subject was aware of this and would be expected to move his target as soon as the lens was changed. It might be thought that this would reduce the second latencies. That it did not do so called for special investigation.

Procedure

Thirteen subjects who were still available, and who had mean latencies > 5 sec in Expt I, were re-examined 5 or 6 months later and under conditions which were similar in general procedure to the above, but which differed in the following respects: (1) The two cylinders used were of equal diameter (1·9 cm) but unequal height (2 m and 1·8 m). (2) Just before the lens was to be introduced, the experimenter displaced his trolley so that the targets were no longer apparently equidistant from the subject. The displacement was to a constant distance equal to approximately the mean maximum displacement achieved by all previous subjects. In metric terms this was 41·4 cm at a viewing distance of 3·5 m. This displacement was made either in the direction in which the subject would be expected to move *his* trolley in response to the altered disparity (referred to as PRO) or in the opposite direction (CON). The reasoning was that the subject, knowing that the targets were out of alignment, would, *as instructed*, start to correct this as soon as the lens was introduced. In both cases movement would be initiated in one case in the direction of the expected final position (PRO) and in the other in the opposite direction.

Each subject was tested twice with the lens over the same eye. On one occasion the induced movement was PRO and on the other CON. Half of the subjects experienced the order PRO/CON. The mode of time intervals between sessions was 7 days with a range of from 1 to 47 days.

Table 5. *Mean latency differences between conditions and experiments*

	$\bar{D}$ (sec)	t*	D.F.
Expt II			
Session 1 *vs.* 2	2·8 (1 > 2)	0·95	11† (N.S.)
PRO *vs.* CON	4·7 (P > C)	1·69	11 (N.S.)
Expt I *vs.* II (first latency)	8·2 (1 > 2)	1·12	12 (N.S.)

* t-test for related means.

† One subject whose latency exceeded 2 min in session 2, Expt II, was omitted.

Results

The mean latency differences between the two conditions, the two sessions and between Expts I and II are shown in Table 5. It would appear that the induced movement, released as soon as the lens was introduced, had only a limited effect on latency.

In addition, of the eight subjects used in this experiment who had exhibited reversals in Expt I, five behaved in this experiment as if they were under the influence of the same tendency, in both sessions: in the CON condition they proceeded far enough beyond the point of target equidistance to satisfy the original criteria; in the PRO condition they either made no move at first or made a move in the demanded direction and then reversed it, also to criteria. All subjects who did not exhibit reversals in the PRO condition made an initial adjustment to the vicinity of target equidistance, paused there for more than 5 sec and then continued in the predicted direction.

It may therefore be concluded that the results obtained in Expt I on the change of lens represent reliable tendencies that have to do with latency and not with some form of inattention.

Discussion

The above findings largely confirm those previously reported. Latencies occur and exhibit large inter-person differences which have some reliability under the present conditions. There appear to be three main classes of response: a response in the predicted direction with a short latency (< 5 sec); a response in the predicted direction with a longer latency; and a response which is initially in the 'wrong' direction. These classes will be referred to, respectively, as no latency (NL), simple latency (SL), and reversals (*Rev*). There was some evidence from subjective reports that a fourth category may exist, that of oscillatory response (*Osc*).

If this classification is reliable, the use of latency as an index of other personal characteristics would have to take it into consideration. The use made of the duration of latency by such authors as Gardner *et al.* (1959) and Martin (1954) would have confounded *Rev* with SL. Ames' (1946) explanation of latency in terms of the knowledge the subject has of the real situation is clearly invalidated by the results of Expt II.

Ittelson's (1960) report that many subjects achieve a full response to the effects of the lenses by successive approximations is amply confirmed. This tendency is reminiscent of Woodworth's (1947) description of perception as a process of 'trial-and-check, trial-and-check' (p. 123), and was also reported by Wheatstone (1852) with the pseudoscope. However, the causes of latency and of these successive approximations have yet to be established.

Reversals require some discussion. Woodworth & Schlosberg (1955) reported reversals in what they call the 'postage stamp experiment' (p. 477) in which an induced convergence occasions a false fusion. The fused figure is usually perceived as closer than the real one and smaller. However: 'The apparent distance is not so definite; sometimes the combined figure appears nearer, sometimes farther than the objective distance.' They went on to suggest: 'When the apparent distance is less than the real distance, convergence is apparently the cue; but when the figure appears farther away than it is, the apparent size is the dominant cue.' They attributed this changed apparent size to the workings of size–distance invariance and since they supposed the convergence cue to be primary and the size change to be a derivative of it, they were placed in a paradoxical situation when the observer responded to the size change only. This finding, together with others, led to their suggestion that perception is an activity which functions at several levels.

A way out of their paradox may be found in a suggestion made by Gillam (1964). If the geometrical optics outlined in the introduction to this paper are accepted as a reasonable representation of the source of information available in binocular depth resolutions, we may also conclude that the same proximal stimulus configurations provide information for the achievement of *binocular shape resolutions*. Gillam, in making this point, states: '...since binocular retinal stimulation is a cue to spatial localization in general it is not only a cue to slant, but also a cue to shape. It can be shown that the shape distortion induced by aniseikonic lenses can be predicted directly by means of the optical model from the binocular stimulation present when aniseikonic lenses are worn. Therefore, it is not necessary to postulate invariance as an explanation of the size–shape distortion' (pp. 150–1).

The geometry of the supposed event, which we shall refer to as *reversed perspective*, is depicted in Fig. 4. It had already been described by Ogle (1950) and the notion of a binocular shape resolution had also been worked out by Ames and his colleagues and had been incorporated into the binocular distorted room (Kilpatrick, 1952, pp. 50–2).

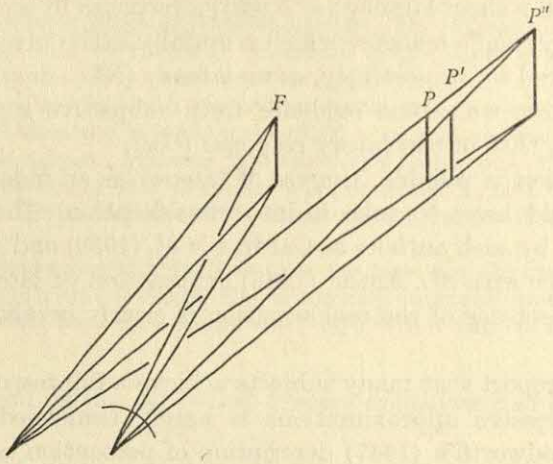


Fig. 4. The geometry of reversed perspective when a lens is worn over the right eye. The apparent location and size (and, therefore, shape) of P are altered.

The paradox is avoided by assuming that the two resolutions are parallel but independent primary events, normally concordant and indistinguishable but becoming discordant and distinguishable under the artificial conditions of the various experiments cited: in the case of the postage stamps, the object is not seen as smaller *because* it is seen as nearer but is seen both as nearer and smaller because of the proximal binocular information. Reversals may then be due either to a shorter latency for shape resolution than that for depth resolution, and to some subjects responding to the apparent shape change, or to a reliance on shape as an index of orientation, when oscillations would be due to a vacillation between the two resolutions.

Some phenomenological evidence supports this interpretation. A number of adults questioned have reported apparent reversed slants in natural objects viewed through a lens when the viewing distance has been more than a few yards, while the same objects take on a predicted slant at shorter distances. This is contrary to geometrical prediction (cf. Ogle, 1950, p. 162) and to Ames' observations (1946, pp. 343, 356). This seems to indicate an increasing reliance of shape as an index of orientation and as a function of increasing distance, as Gillam (1967) has verified.

It is interesting that previous authors have not noticed reversals with aniseikonic lenses. This may be due to a reliance on subjective report rather than on measurement, or to the transitory nature of the event: it always gave way eventually, in the present instance, to a perceived distortion in the predicted direction.

If the above account of the events leading to reversals is correct, that the subject uses the reversed perspective as an index to orientation, reversals might not occur if the targets did not possess shape in the usual sense of the word. This was examined in the following experiment.

Experiment III

The desired situation was achieved by the use of edges as targets, as shown in Fig. 5.

Apparatus

The targets were two black metal plates mounted on the trolleys. Their inside edges occupied the same position as had the inside edges of the cylinders used in Expt II. The plates were orientated at approximately right angles to the line of regard when it was directed to each in turn.

The subject looked through two rectangular apertures in a screen placed between him and the lens. An adjustable septum was mounted on the rear of this screen. The second screen (see Fig. 5) was placed at approximately 5 cm from the eye. The subject was unable to see the outside, top or bottom edges of the plates. At the distances involved, the frame within which his view of the targets fell was blurred. The blurred area was narrow and the subject obtained a clear view of the target surfaces subtending a visual angle of approximately 21° in the horizontal and 7° in the vertical; of this horizontal angle, 13° was taken up by approximately equal extents of two targets, 8° by the gap between them. At the viewing distances (see Expt I) the surface texture of the plates apparently could not be detected. A green fabric chalkboard, which contained some surface marks by which it could be located in space, was used as a background. (When a diffused white background screen was used in exploratory studies, the subject's adjustments of the apparatus tended to become chaotic, even without lenses.)

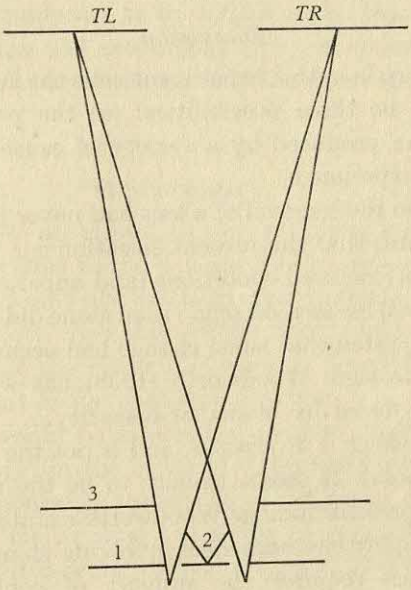


Fig. 5. Plan of viewing conditions in Expt III. 1, Screen carrying the apertures; 2, adjustable septum; 3, distant reduction screen; *TL*, *TR*, binocularly visible sectors of plate targets.

Subjects

Twenty-eight subjects who had taken part in previous experiments involving aniseikonic lenses.

Procedure

There were two sessions separated by about 1 week. All subjects were tested with the lens over each eye on both occasions. For half of the subjects, the order was *R–L* on the first occasion and *L–R* on the second. For the other half the orders were *L–R* . . . *R–L*.

The procedure was: five practice settings; first lens in; latency tracking (2 min); lens out; apparatus resetting by the subject; five practice settings; second lens in; latency tracking (2 min); lens out; apparatus resetting by the subject.

This provided four estimates of latency for each subject, both on the introduction and on the removal of the lens. This last estimate was made because preliminary investigations had indicated that there might be some latency when the lens was removed under these severely reduced conditions and because Vernon (1937, pp. 140–1) had reported similar events.

Results

Only the most important aspects need to be summarized. They were:

(*a*) The reduced conditions decreased the reliability of the subject's latency durations, the rank order (ρ) correlation between the two sessions for each eye separately being: R eye $= +0{\cdot}37$, L eye $= +0{\cdot}79$.

(*b*) The over-all mean latency on the introduction of a lens was 19·53 sec with S.D. 26·7.

(*c*) On twenty-nine occasions in 112 there were latencies in excess of 5 sec on the removal of a lens. This had not occurred in previous experiments and confirms Vernon's (1937) observations. The mean duration of these was 15·31 sec with S.D. 10·36. (Mean duration for all 112 occasions = 6·20 sec; S.D. 7·59.)

(*d*) There were twenty-eight cases of reversal on the introduction of a lens.

Discussion

The occurrence of reversals in this experiment calls the suggested explanation into doubt. There appear to be three possibilities: (*a*) the present experiment is not relevant; (*b*) reversals are produced by a variety of causes; (*c*) shape may still be involved in the present experiment.

(*a*) *and* (*b*). Latencies on the removal of a lens had never been observed in previous experiments. This indicates that the present situation is a very difficult one for the subject: under the severely reduced conditions (and apparently more so if the background is made textureless) his stereoscopic vision alone did not seem to be adequate. Typically, the subject reported that some change had occurred, but he did not seem to be able to identify its sign. Woodworth (1938) has warned against supposing that 'the depth effect is *forced* by binocular disparity.... Binocular disparity is a stimulus, but the depth effect is a *response*, and is not the only response which the organism has at its disposal. It seems usually to be the most satisfying response' (p. 658). In view of the present findings Woodworth's statement might be expanded in the following way: in many circumstances involving changes in existing binocular viewing, the depth effect requires the support of confirming non-stereoscopic information to become effective. Where such information is contradictory (reversed perspectives) or absent (no real movement in the targets), a certain time is required for the depth resolution to become effective and under some circumstances the direction of the depth change may be misinterpreted.

To the extent that the normal functioning of binocular vision is being examined, therefore, the present experiment may be irrelevant and the reversals obtained may be due to causes which are different from those producing the reversals of Expts I and II.

(*c*) It could be argued that the space between the target plates has a shape bounded by the target edges and by the upper edges of the reduction screen. The shape of this space would be altered as a function of alterations in the apparent lengths of the two

inner edges of the plates. Such an explanation would require that the apparent judged lengths of these inner edges are not limited by the upper and lower edges of the reduction screens.

It is not necessary, however, to depend on appearances for an explanation if 'indices to shape change' is substituted for 'apparent shape change' in the explanation. The argument would be that the *response* is made to the proximal stimulus change which is the same in Expt III as it was in Expts I and II: a change in the relations of the images at the two eyes, which change, by our hypothesis, entails *both* a depth and a shape resolution. The reversals found in all the experiments would be due to the same cause: some subjects rely on the resolved shape as an index to orientation. There is no need to refer to conscious experience as an arbiter of response: in Expts I and II there was a concordance between the response and the conscious experience, in Expt III there may not have been. There are countless examples in perception of responses to indices which are not available to consciousness, responses which are sometimes veridical and sometimes erroneous.

In view of the demonstrations of Ames and his colleagues, it is to be expected that some change in the targets *will* be perceived at a conscious level. This will include that one target appears to be farther away than the other, but what compensation and distortions will accompany this perception cannot be predicted with certainty and may well vary between individuals (cf. the rotating trapezoidal window).

Conclusions

Certain important problems are raised. In the first place, and in view of the above argument, it would be interesting to try to separate the two suggested resolutions of orientation and of shape. It would also be interesting to try to separate the effects of shape resolution from any size constancy effects that might accompany the slant resolution: it is possible that the shape distortion is accentuated when the depth resolution is a perceived slant in the predicted direction.

In the second place, it would be interesting to know if reliance on perspective changes represents a primary (in the sense of unlearned) tendency: this might be answered by examining responses to the effects of the lenses as a function of age.

Finally, it may be noted that, in respect of all the phenomena observed, there were indications that important inter-ocular differences exist. These require to be investigated and, if confirmed, would rule out the use of responses to the effects of aniseikonic lenses in personality studies and would present an intriguing problem for any theory of stereoscopic vision.

REFERENCES

AMES, A., JR. (1946). Binocular vision as affected by relations between uniocular stimulus patterns in commonplace environments. *Am. J. Psychol.* **59**, 333–57.

AMES, A., JR., GLIDDON, G. H. & OGLE, K. N. (1932). Lenses for changing the size and shape of dioptric images. *Am. Dist. Serv. Found. Opt.* **1**, 61–70.

GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B. & SPENCE, D. P. (1959). Cognitive control: a study of individual consistencies in cognitive behaviour. *Psychol. Issues* **1**. (Whole no. 4.)

GILLAM, B. J. (1964). Space perception with aniseikonic lenses. (Unpublished doctoral dissertation, Australian National University.)

GILLAM, B. J. (1967). Changes in the direction of induced aniseikonic slant as a function of distance. *Vis. Res.* **7**, 777–84.

ITTELSON, W. H. (1960). *Visual Space Perception*. New York: Springer.

KILPATRICK, F. P. (ed.) (1952). In *The Ames Demonstrations in Perception*. Princeton, N.J.: Princeton University Press.

LANCASTER, W. B. (1938). Aniseikonia. *Arch. ophthal.* **20**, 907–12.

MARTIN, B. (1954). Intolerance of ambiguity in interpersonal and perceptual behaviour. *J. Pers.* **22**, 494–503.

OGLE, K. N. (1950). *Researches in Binocular Vision*. Philadelphia: Saunders.

OGLE, K. N. (1962). The optical space sense. In *The Eye*, vol. IV, pt. II, ed. H. Davson. New York: Academic Press.

VERNON, M. D. (1937). The perception of distance. *Br. J. Psychol.* **28**, 1–11; 115–49.

WHEATSTONE, C. (1852). On some remarkable and hitherto unobserved phenomena of binocular vision: II. *Phil. Mag. Ser.* **4** (3), 504–23.

WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

WOODWORTH, R. S. (1947). Reinforcement of perception. *Am. J. Psychol.* **60**, 119–24.

WOODWORTH, R. S. & SCHLOSBERG, H. (1955). *Experimental Psychology*. New York: Holt.

(*Manuscript received* 20 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 351–356
Printed in Great Britain

STEREOSCOPIC VISION AND ANISEIKONIC LENSES. II

By G. N. SEAGRIM
Australian National University, Canberra

Two experiments are reported which examine some implications of earlier experiments (Seagrim, 1967). It is found that latency in the response to the effect of aniseikonic lenses persists under altered conditions but is radically reduced when an occluder is placed before either eye (but not both) for 10 sec while the lens is being placed in position. It is also found that latency is positively related to age.

In a previous report (Seagrim, 1967) details were given of the behaviour of subjects wearing an aniseikonic lens in the 90° position over one eye. In particular, measurements were made of the delay, or latency, of the predicted response. In Expt II it was found that latencies persisted even when the subject was required to re-align the target as soon as the lens was introduced. This finding ruled out the possibility that he was inattentive and made it unlikely that his binocular vision had lapsed into temporary disuse. It also seems to rule out the possibility that those subjects who failed to respond immediately did so *because* of a tendency to perseverate with the previously established perceptual interpretation of the situation (see Woodworth, 1938, p. 680).

The apparatus then used had two undesirable features. First, the viewing distance was rather great for binocular vision (260–400 cm). Secondly, the subject was required to displace one of the objects whose visual location was being affected by the lens. Ames (1946) has claimed that the effects of the lens tend to disappear when the wearer moves and it is possible that the same thing happens when the target moves. Furthermore, because of the probable intervention of size–distance effects (see Seagrim, 1967), perceptual distortions would accompany any movement of the targets and these distortions might affect the subject's interpretation of the information.

Method

In order to overcome these difficulties, a new apparatus was designed.

Apparatus

The viewing distance was reduced to 164 cm. The target consisted of two vertical rods, 4 mm in diameter, suspended from a cross-bar. The distance between the rods was 10 cm, representing a visual angle of 3° 30′.

A tracking beam made of 'Meccano', ⅃ in section and 31·5 cm long, was mounted horizontally, 6 cm (2° 6′) below the lower extremities of the target rods.

The lower one-third of an aniseikonic lens was sliced off and the remainder was positioned so that the subject viewed the targets through it and the tracking beam below it (see Plate 1): he thus obtained an uninterrupted binocular view of the tracking beam.

The lenses were mounted on stalks which were hinged by means of a ball-and-socket joint. This allowed flexibility in their positioning. (If the lens is rotated around its vertical axis the apparent distance of the target is altered. Lenses were so positioned that the targets and the beam appeared to be at the same distance from the observer.)

The lenses were suspended from a sliding bar which could be moved across the subject's front by the experimenter. In this way it was possible to have a lens in front of each eye, a lens in front of either eye, or no lens in front of either eye.

The target could be rotated by the experimenter and the tracking beam by the subject, who was required to keep it parallel to the target rods at all times. All displacements of the targets and of the beams were automatically recorded against time and the introduction or removal of a lens in front of the subject's eye was recorded by an event marker.

The subject's head was supported in a chin rest. The height of this was adjusted so that the bottom edge of the truncated lens lay in a line between the subject's pupil and the gap between the targets and the beam. This alignment was made with a non-magnifying (dummy) lens of the same dimensions as the lens to be used.

The subject's view of the targets was restricted by two reduction screens. The lower of these consisted of a horizontal rod, 1 cm in diameter, placed 28 cm from the subject's eye. Its height was such that, when the subject was correctly positioned, its blurred image obscured his view of the lower extremities of the targets in whatever position they might be. It thus coincided with the line of the lower edge of the lens. The upper reduction screen consisted of a black card, whose lower horizontal edge was 66 cm from the subject's eye and obscured his view of the upper parts of the target apparatus. As a result, the subject obtained a view of the middle segments of the targets, subtending a visual angle of 6° 18′ (18 cm). In other respects the subject's field of vision was not restricted. Illumination was by overhead neon lighting and by a diffused background illumination located 25 cm behind the targets. This was provided by a milk-white 'Acralite' dome, 36 in. in diameter, behind which two incandescent lights were directed at the room wall. The over-all result was shadowless and the targets appeared as two dark rods against a white background of indeterminate distance. Extraneous cues to the direction of the movement of the targets were shielded from the subject's view.

The subject's view of the apparatus is shown in Plate 1 and general views in Plate 2 (*a*) and (*b*).

EXPERIMENT IV

The purpose of the experiment was to examine latencies under the present conditions. As it had been found, in informal observations, that latencies were reduced if the subject was taken with his eyes shut to the scene to be viewed through a lens, it was decided to examine the effects on latencies of the introduction of an occluder between the eyes and the target while the lens was being placed in position.

Procedure

A 4 % lens was used in the 90° position before the left eye. An occluder of sanded Perspex was added to the apparatus. It could be interposed between the lens and either or both eyes.

Each subject underwent two trials separated by not less than one week. The first trial constituted the familiarization procedure and its results were used as a standard against which performance on trial 2 could be assessed.

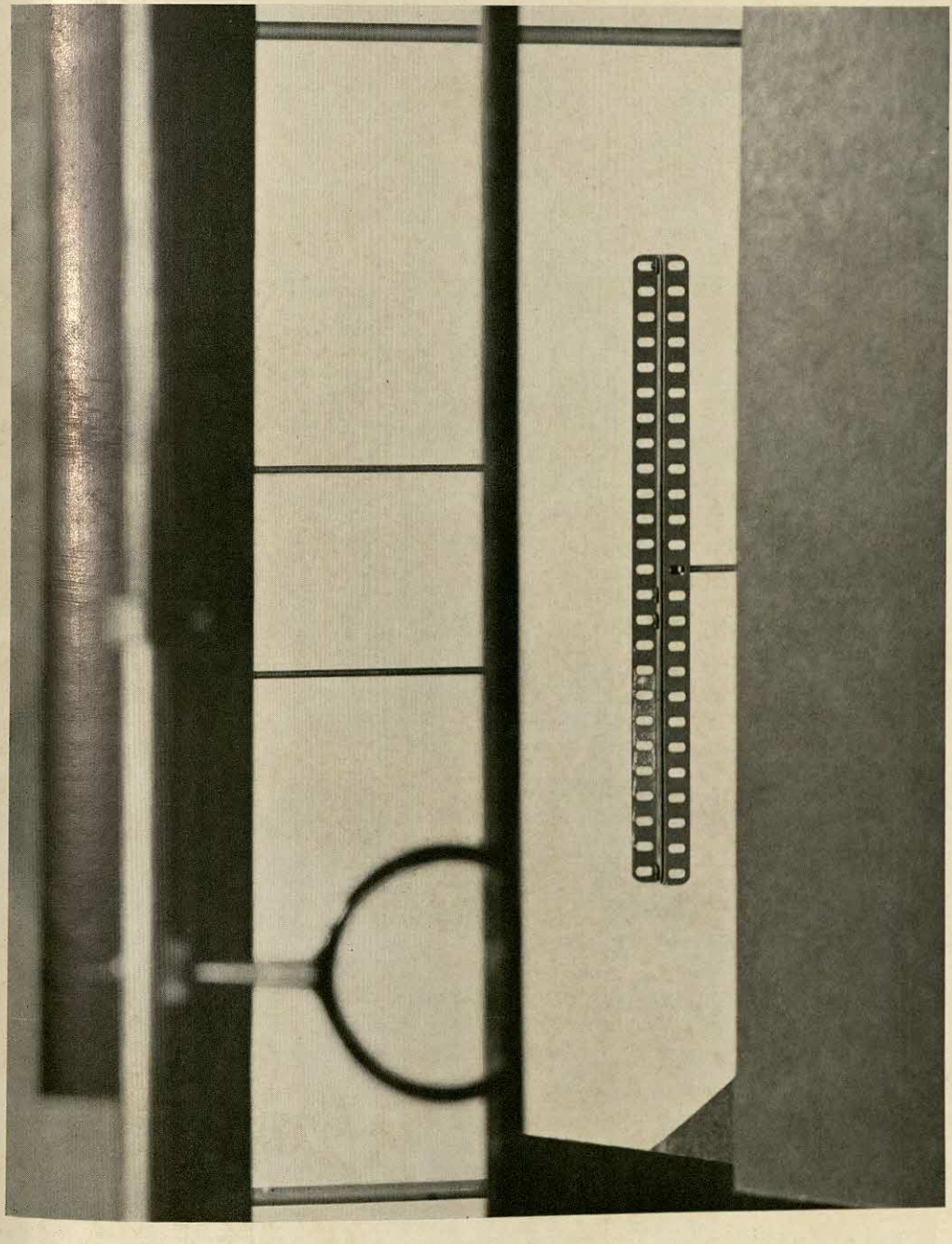
Trial 1. Five practice settings (see Seagrim, 1967, for procedural details) with a dummy lens over the left eye; 10 sec pause; dummy lens out; left lens in; latency tracking (2 min); lens out.

Trial 2. The subjects were allocated to one of four groups in the order of their attendance for trial 2. The four groups were: C, control; B, occluder over both eyes; R, occluder over right eye; L, occluder over left eye. For group C the procedure was the same as for trial 1. For groups B, R and L practice settings with occluder were interposed each time before the experimenter readjusted the targets (the occluder was removed after a total time of 10 sec). The sequence was: dummy lens out; occluder in; lens in; occluder out (10 sec after introduction); latency tracking (2 min); occluder in; lens out; occluder out; apparatus readjusted by the subject.

Before trial 2, all subjects were warned that the procedure would be approximately the same as in trial 1 but that they were not to assume that their experience would be the same; they were to respond as the situation demanded and were not to try to anticipate. Subjects in groups B, R and L were shown how the occluder would be used. Fixation was not demanded.

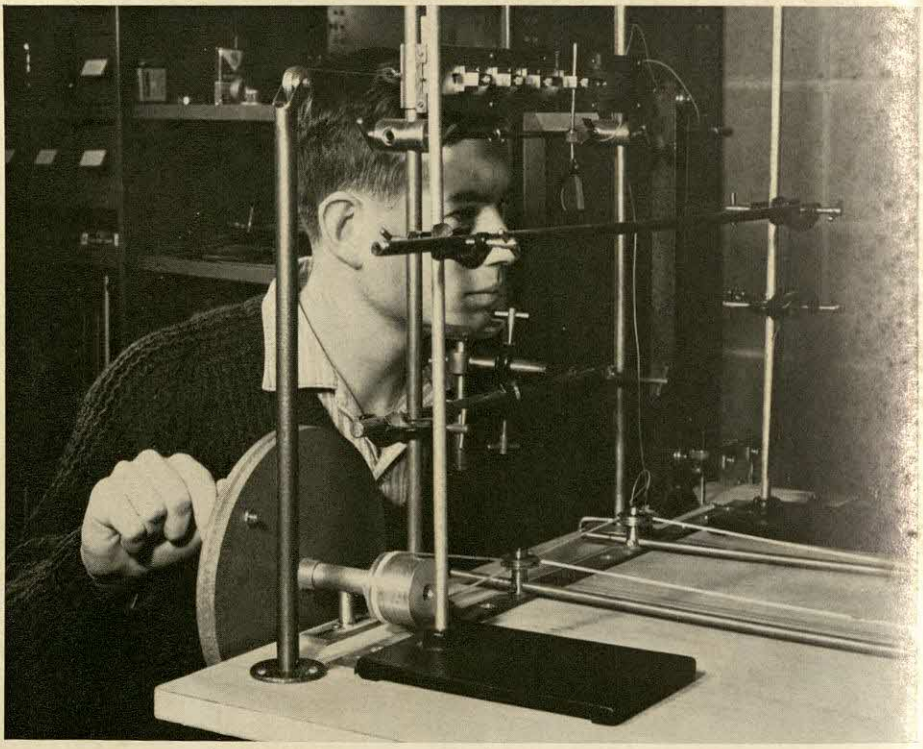
Subjects

Results were obtained from fifty-nine young adults who had not previously experienced the effects of the lenses. (Sixty-five subjects underwent trial 1, but six did not reappear for trial 2.)

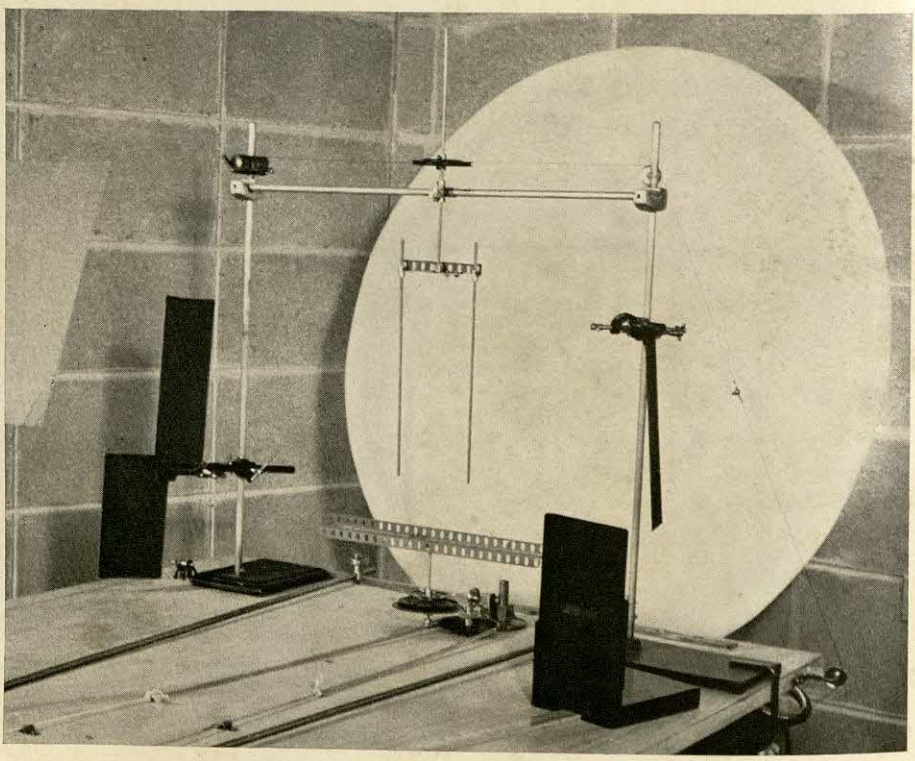


G. N. SEAGRIM

(*Facing p.* 352)



a



b

G. N. SEAGRIM

Results

The results are set out in Tables 1 and 2. The latencies of two of the fifty-nine subjects exceeded 2 min on both occasions; for this reason their results have not been included in the tables. One of these subjects was in group B and the other in group R, hence the final sample numbers were: group C, 15; group B, 13; group R, 14; and group L, 15.

Table 1. *Frequency of reversals by groups and of oscillations by trials*

	Reversals	
	Trial 1	Trial 2
Group C	1	—
B	2	2
R	3	—
L	—	—
Oscillations	7	10

Table 2. *Comparison of durations of latencies (sec) by groups and by trials*

Group ...	C	B	R	L
n ...	15	13	14	15
Trial 1				
Mean	23·7	39·1	32·6	23·4
S.D.	16·5	42·6	24·2	20·7
Trial 2				
Mean	17·5	23·5	4·4	4·4
S.D.	15·44	32·3	5·9	7·4
Reduction (%)	26·3	39·9	86·5	81·2
t (trial 1 − trial 2)	1·1	1·1	4·23**	4·0**

** $P < 0{\cdot}001$.

Discussion

While the frequency of reversals was reduced in all conditions by comparison with Expts I and II, oscillations now appeared in the tracking record. These oscillations, already reported by other authors (see Seagrim, 1967), may represent continuing tendencies towards reversals under conditions in which the indices leading to reversals are relatively weakened: there may be a greater initial reliance on the depth resolutions at the shorter viewing distance. However, as the reduction in the frequency of reversals and the appearance of oscillations cannot be attributed with any certainty to any one of the conditions which were changed in this experiment (distance, tracking device, target size, lens magnification), they will not be discussed any further at this juncture.

The latencies without the occluder (trial 1) and in trial 2 for group C are about the same as they were under the conditions of Expts I and II (Seagrim, 1967), which indicates that latency is a general phenomenon and not specific to certain conditions of viewing which are unfavourable to binocular vision. It is a condition into which some subjects have fallen and which is not corrected by measures which would seem to ensure that the subject is using his binocular vision: the alerting

procedures of Expt I, the forced movements of Expt II, the edges of Expt III and the use of a binocular occluder in the present experiment. That the condition is not irreversible is shown by the dramatic effects of using a monocular occluder.

Table 3. *Distribution of latencies obtained in Piaget's laboratory*

Latency (sec)	Age of subjects: 5–6 yr (n = 30)	Age of subjects: Adult (n = 18)
0–9	16	5
10–19	10	6
20–29	2	1
30–39	—	1
40–49	2	3
50–59	—	—
60–69	—	2

It seems to be very unlikely that latency is associated with defects in vision. There was no relationship between the subjects' known eye condition and latency nor between measured stereoscopic acuity and latency except in the case of those very rare subjects who were functionally monocular because of an earlier strabismus: they were insensitive to the effects of the lens under any circumstances and were not used as subjects. All other subjects with known visual defects wore their correcting spectacles during the experiments and, as a group, did not differ, in respect of latency, from those who did not have known visual defects.

There seems to be no way of accounting at present for the differential effect of monocular as opposed to binocular occlusion. The difference did not appear in another experiment (unpublished) using the apparatus of Expt III (Seagrim, 1967). In that case latencies were reduced under all three conditions of occlusion taken together but not for each separately; the greatest reductions were obtained under the condition of binocular occlusion. It is possible that group B of the present experiment was atypical, as is suggessted by the relatively large S.D.S obtained.

It now seems fairly safe to conclude that the tendency for some subjects to respond slowly to the effects of aniseikonic lenses represents some form of acquired tendency, what Segall, Campbell & Herskovits (1963) refer to as 'a habit of perceptual inference'.

If this is so, latency should increase as a function of age. Our own informal observations with children in natural viewing conditions suggested that latencies were very short. This possibility received support from Piaget's laboratory in Geneva in 1963, where Dr Marianne Denis obtained the distributions of latencies shown in Table 3 (personal communication; reproduced with the permission of Professor Piaget). The prediction was further examined in Expt V.

EXPERIMENT V

Apparatus

As in Expt IV but without the occluder and without the bar which obstructed the lower extremities of the targets (because of the difficulty of positioning young subjects; its removal adds to possible contradictory indices and would therefore tend to increase latencies if the present reasoning is correct).

Subjects

Young males of the numbers and ages given in Table 4.

Procedure

All subjects completed two trials; half of them with the lens on the right eye first and half with the lens on the left eye first. Trials were separated by not less than 5 min, during which the subject performed a perceptual test not involving the use of lenses. Each trial consisted of: practice settings; lens in; latency tracking (2 min); lens out.

Results

The results are set out in Tables 4 and 5. One of the subjects tested has been excluded because his performance could not be scored.

Table 4. *First and second latencies (sec) by age-groups*

Age (years, months)											
16.5–18.6		12.0–13.7		11.2–11.11		10.4–10.9		9.0–9.11		8.2–8.10	
1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
75	86	14	27	6	7	3	30	19	5	5	5
56	26	62	103	41	120	11	19	32	120	8	6
9	5	14	5	29	18	2	2	8	3	6	4
21	5	5	38	13	99	16	3	3	4	8	4
39	113	37	40	11	11	22	17	3	4	27	6
99	20	26	68	76	33	25	7	4	17	5	5
120	45	45	7	53	14	—	—	72	5	2	3
84	32	8	7	2	3	—	—	7	5	—	—
—	—	22	73	—	—	—	—	13	11	—	—
—	—	20	13	—	—	—	—	7	2	—	—
Mean 62·9	41·5	25·3	38·1	28·9	38·1	13·2	13·3	16·8	17·6	8·7	4·7

Table 5. *Distribution of reversals and oscillations by ages and by trials*

	Reversals		Oscillations	
Age	T_1	T_2	T_1	T_2
16.5–18.6	—	2	—	—
12.0–13.7	2	3	1	2
11.2–11.11	—	2	1	2
10.4–10.9	—	1	—	1
9.0–9.11	1	—	—	2
8.2–8.10	—	—	1	2

Discussion

The increase of latency with age is clear and appears to exhibit a sharp change between the ages of 9 and 10 years. This supports the notion that latency represents an acquired or maturing tendency to neglect changes in strictly stereoscopic indices to depth. It is suggested that this tendency need not be regarded as inefficient or retrograde but that it may, under our conditions of life, represent an efficient adaptation: the binocular attention can be freed from the need constantly to monitor changes in depth relationships if other simpler indices to those changes exist and are reliable.

The increase in reversals with age and, in this case, of reversals and oscillations with practice, suggests that the same basic mechanisms may be involved: a reliance on, and in this case, a response to, non-stereoscopic indices to depth.

Finally, this increasing reliance on, or exploration of, alternative solutions suggests the operation of some factor which might be related to intelligence and to occupation. In this connexion Piaget's notion of *perceptual activities* (1961) suggests itself and an essay in this direction has been attempted elsewhere (Seagrim, 1966).

Acknowledgement is made to Mr P. E. McKeown, Headmaster, Canberra Grammar School, for permission to conduct Expt V at his school.

REFERENCES

AMES, A. JR. (1946). Binocular vision as affected by relations between uniocular stimulus patterns in commonplace environments. *Am. J. Psychol.* **59**, 333–57.

PIAGET, J. (1961). *Les Mécanismes perceptifs.* Paris: Presses Universitaires de France.

SEAGRIM, G. N. (1966). L'exploration perceptive en vision binoculaire. *Sci. Comport.* **1**, 271–8.

SEAGRIM, G. N. (1967). Stereoscopic vision and aniseikonic lenses: I. *Br. J. Psychol.* **58**, 337–50.

SEGALL, M. H., CAMPBELL, D. T. & HERSKOVITS, M. J. (1963). Cultural differences in the perception of geometric illusions. *Science, N.Y.* **139**, 769–71.

WOODWORTH, R. S. (1938). *Experimental Psychology.* New York: Holt.

(*Manuscript received* 20 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 357–364
Printed in Great Britain

EFFECTS OF A NOVEL SITUATION AND OF ANXIETY ON TWO GROUPS OF DEPENDENCY BEHAVIOURS

By MIRIAM K. ROSENTHAL
Birkbeck College, London

The habituation of two groups of dependency behaviours in a novel situation was investigated under two levels of anxiety provocation. The two groups of dependency behaviours were 'attention seeking' and 'proximity seeking'. Sixty-four girls aged 3–5 were assigned to two experimental conditions: high and low anxiety. Each child was seen twice and each session lasted 30 min. Frequency of dependency behaviour was recorded by an observer according to previously defined behaviour categories.

The principal findings were: (1) under high anxiety conditions frequency of attention seeking decreases with time (habituation of curiosity) and frequency of proximity seeking increases when there is a gradual increase in number of fear-provoking elements; (2) under low anxiety conditions frequency of attention seeking decreases with time while only a negligible decrease of proximity seeking takes place; (3) long-term habituation, over 3 or 4 weeks, does not occur for either kind of behaviour; (4) an unexpected interaction indicated that attention seeking decreases significantly faster under conditions of low anxiety than under conditions of high and sustained anxiety, although no main effect of anxiety on attention seeking was found. The results are interpreted in terms of elements in the situation provoking curiosity or anxiety.

The effect of 'novelty' on behaviour has been investigated extensively in the last decade (Berlyne, 1960; Fiske & Maddi, 1961). Different degrees of novelty have been known to arouse both avoidance and approach behaviour. Most theories which have tried to incorporate and explain these effects have utilized concepts such as 'incongruity-dissonance' or 'arousal of incompatible cognitive processes' (Hebb, 1946, 1949; Hunt, 1960) or 'optimal arousal level' (Fiske & Maddi, 1961). The basic hypothesis concerning these different effects of novelty is that some unfamiliar elements in the novel situation, most probably those that are incongruous with expectations, arouse fear and alarm in the organism, while the more congruous elements, which are still novel, arouse 'curiosity'.

For an empirical approach, the most obvious question to ask is how can the novel situation be manipulated so as to cause a shift from avoidance to approach or vice versa. Some shifts of this nature have been reported; those of most relevance to the present investigation have been in a 'social' context. Thus, Harlow's work with monkeys demonstrated that temporary contact with a mother or surrogate mother affected exploratory (i.e. approach) behaviour (Harlow, 1953). Arsenian (1943) demonstrated that the presence of the mother with the child in a strange room tended to have 'security value' in that symptoms of fearful reaction to the novel situation decreased. Sarason *et al.* (1960) suggested that the anxious child is cautious, with a tendency to set up protective barriers against stimuli. If one could reduce anxiety level presumably more exploratory behaviour would be found.

The main starting-point for the present study was the suggestion made by Harlow (1953) that some form of dependency behaviour toward a mother might reduce the threatening aspects of unfamiliar, novel situations, thus allowing exploratory behaviour to take place. Following this suggestion it may be argued that by exposing the organism (monkey or child) to an unfamiliar and somewhat threatening situation

constructed in such a way as to allow dependency behaviour to take place, the organism will exhibit more dependency behaviour the more threatening the unfamiliar situation is.

Taking into account the phenomenon of stimulus satiation (Glanzer, 1958) or habituation (Hinde, 1961) which suggests that reaction to novelty decreases with continuous exposure to a novel situation, it is to be expected in addition that the organism (monkey or child) will exhibit more dependency behaviour upon immediate exposure to an unfamiliar situation than it will later on.

In sum, we should expect a young child who is exposed in the presence of an adult to an unfamiliar situation which is fear-provoking (owing to a conflict between familiar and unfamiliar elements) to show a high frequency of dependency behaviour which gradually decreases as the child habituates to the situation.

A study which manipulated anxiety level experimentally and used various groups of dependency behaviours has shown that an increase in anxiety level will bring about an increase in the frequency of some dependency behaviours and not in others (Rosenthal, 1967). In that study the group of dependency behaviours affected by anxiety level was classified as 'proximity seeking' (e.g. 'being near', 'clinging'). The group which did not seem to be affected by anxiety was classified as 'attention seeking' (including asking for approval and help). In regard to the reaction to novelty, the threatening elements in the situation would be expected to provoke 'proximity-seeking' behaviour while the curiosity-arousing elements would be expected to bring about 'attention-seeking' behaviour, expressed by calling the adult's attention to objects and events in the novel situation. Hence the frequency of both groups of responses would be expected to decrease with continuous exposure—one due to habituation of fear and the other due to habituation of curiosity. If a deliberate increase of anxiety-provoking elements in the situation were introduced, however, the two groups of responses would be expected to behave differently. The frequency of 'attention seeking' should decrease with time whatever the number of anxiety-provoking elements, since it is not affected by anxiety but only by novelty. The frequency of 'proximity seeking', however, should decrease when anxiety level is relatively low, but increase as additional anxiety-provoking elements are introduced into the situation.

Long-term habituation is not to be expected: a child who is confronted with an unfamiliar situation once and then confronted with it again several weeks later cannot be expected necessarily to show less dependency behaviour on the second confrontation. Such a longer-term habituation might be expected for two confrontations which were closer together in time.

METHOD

Subjects

The subjects were sixty-four nursery-school girls aged 3–5 years. Most of them came from Inner London Education Authority nursery schools and they and their mothers participated willingly.

Procedure

(*a*) *Experimental observation*

The experimental room was an observation room with a one-way screen located in an office building. In the room were toys, a stationery cupboard, a desk, two adult-size chairs and two child-size tables and chairs. A flowered curtain partially divided the room. The walls were

decorated with coloured papers and drawings of faces. The toys consisted of an open-roof dolls' house with furniture, dolls, 'Lego', puzzles, toy telephone, paper and crayons, magazine pictures and scissors, beads with strings, blackboard and coloured chalk, 'mother's' dress, doll's pram and cradle with four dolls, doll's swing, small cars and mechanically moving animals.

The experimenter (female) met the child in the nursery school and escorted her to the observation room, being quite friendly on the way. Once there the experimenter excused herself ('to repair some toys') leaving the child in the room with her mother or with another woman (mother's surrogate). The experimenter came in again at the end of the session (30 min) and took the child back to the nursery shool. While the child was in the room the experimenter observed and scored the behaviour of the child from the adjacent room. A microphone in the observation room, fixed above the adult's desk, enabled the experimenter to hear everything from the observation room. Besides recording the child's behaviour, the experimenter also controlled a tape-recorder which played a pre-recorded sound track, the content of which depended on the anxiety-group to which the child was assigned. The adult in the room was kept busy sitting at the desk behind the curtain and was instructed not to initiate any interaction with the child and to respond to the child's initiation as shortly as possible without actually rejecting the child.

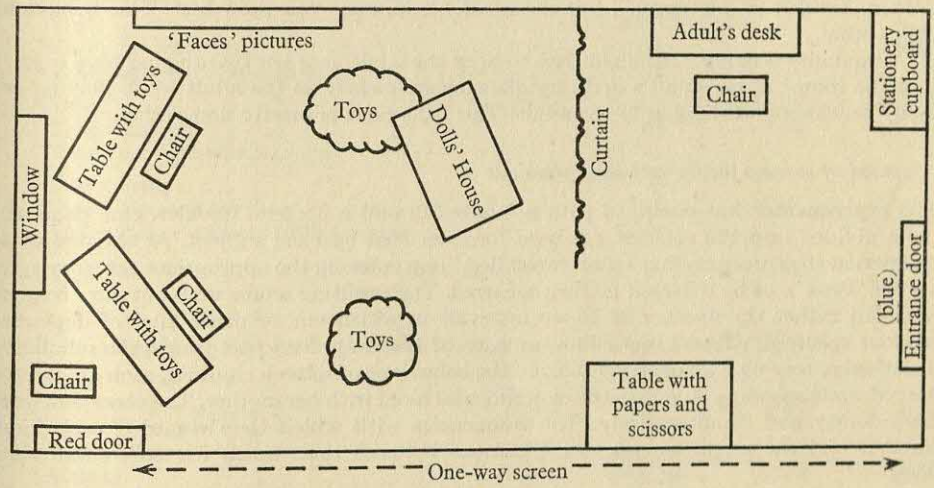


Fig. 1. A schematic plan of the observation room.

(*b*) *Independent variables*

There were three independent variables.

(1) *Anxiety*. Two levels of anxiety were generated experimentally. In the *low-anxiety* condition the child played in the observation room with the various toys available; a picture of smiling faces was hanging next to the toys. The only stimulation from outside the room was the sound of a tape-recorder in the adjacent room, which played first lines of various children's songs at regular intervals of about 3 min. This condition was effectively one of minimal or 'no anxiety', thus functioning as a control for the high-anxiety condition. For the *high-anxiety* condition, the objective was to induce anxiety in a way which was not contaminated by any 'social deprivation' or 'rejection' (see Walters & Ray, 1960), but to generate a state of vague apprehension. The child, on entering the room, faced a slow-burning alcohol lamp standing on a stainless steel tray. Next to it was a pair of scissors, a white paper tissue, and a pencil. The pictures of the smiling faces were replaced with a group of sad faces. The sound-track heard from the adjacent room was made up of the following sounds: a loud banging on a metal object, a child crying, and a high-pitched shriek. Each sound was on for about 20 sec, at regular intervals of about 3 min. The tray with the burning lamp stood on a chair next to the red door leading to the room from which the sounds came. In addition to the taped sound-track, after about 12 min of observation and following a loud continuous shriek, the red door opened very slowly (the

experimenter waited until the child was looking in that direction) and a hand in an arm-length black glove reached slowly in, put out the lamp and withdrew, closing the door once more. Within 2 or 3 min a crying sound was heard.

(2) *Visit.* Every subject was seen twice, once with her mother and once with a 'surrogate', another woman who was very friendly to the child. Hence there were two sets of scores for each subject—one from her first visit (Visit-1) and the other from the second visit (Visit-2). In most cases the second visit followed a few weeks after the first. The order was balanced; thirty-two subjects were seen first with the 'surrogate' and thirty-two subjects were seen first with their mother.

(3) *Time.* This variable refers to time during the 30 min observation session. The first 15 min (Time-1) were distinguished from the last 15 min (Time-2) so that the frequencies of dependency behaviours during the two halves of each observation session could be compared.

(c) *Dependent variables*

There were two dependent variables.

(1) 'Attention seeking' includes behaviour which appears to be employed by the child to get: (i) the attention of the adult, e.g. 'look', 'you know what...', (ii) the adult's praise or approval, e.g. 'look, I did it all by myself', (iii) the adult's help, e.g. 'you show me'. This behaviour is mostly verbal.

(2) 'Proximity seeking': the child tries to keep the adult near her by bringing toys from 'her side of the room' to the adult's desk, by glancing frequently at the adult while playing or by actually touching and clinging to the adult. This behaviour is mostly non-verbal.

(d) *Method of scoring the dependency behaviour*

The experimenter was equipped with a stopwatch and a 3×5 in. booklet. One page of the booklet divided into 120 squares was used for each visit by each subject. At the end of each 15 sec period the corresponding square was filled in by entering the appropriate category symbol ('Att' or 'Prox'), or by a dash if neither occurred. The resulting scores were not pure frequency scores, but rather the number of 15 sec intervals in which one or other kind of dependency behaviour occurred, with an upper limit on each of 120. In order to determine the reliability of the method of scoring, two observers scored the behaviour of sixteen children, each of whom was observed while spending 30 min in the experimental room with her mother. The observers worked independently and simultaneously. The frequencies with which they recorded each kind of behaviour and the product-moment correlations between their totals for each child were as follows:

	f (Observer-1)	f (Observer-2)	r
Attention seeking	326	345	+0·991
Proximity seeking	198	211	+0·990

Hence inter-observer reliability was high.

(e) *Experimental plan*

The plan of the experiment is summarized in Table 1.

Table 1. *Plan of experiment and sample numbers*

Independent variables		(1) Level of anxiety			
(2) Visit	(3) Time	Low (minimal)	High	Factor	n
Visit-1	Time-1	32	32	Visit-1	64
	Time-2			Visit-2	64
Visit-2	Time-1			Time-1	64
	Time-2			Time-2	64
Number of subjects		32	32	64	

Results

The data were submitted to analysis of variance comparing three groups of means: Low anxiety versus High anxiety, Visit-1 versus Visit-2 and Time-1 versus Time-2, as well as the interactions of Anxiety × Visit, Anxiety × Time, Visit × Time and Anxiety × Visit × Time. The analysis was carried out separately for 'attention seeking' and for 'proximity seeking', yielding Tables 2 and 3.

Table 2. *Analysis of variance of frequency of attention seeking*

Source	D.F.	Mean squares	*F*	*P*
Between subjects	63			
Anxiety	1	23·77		N.S.
Subjects within group	62	164·87		
Within subjects	255			
Visit	1	129·39		N.S.
Visit × Anxiety	1	68·05		N.S.
Visit × subjects within group	83	81·40		
Time	1	272·25	13·08	< 0·005*
Time × Anxiety	1	97·50	7·69	< 0·025*
Time × subjects within group	83	20·81		
Visit × Time	1	47·27		N.S.
Visit × Time × Anxiety	1	7·55		N.S.
Visit × Time × subjects within groups	83	23·04		

* One-tail.

Table 3. *Analysis of variance of frequency of proximity seeking*

Source	D.F.	Mean squares	*F*	*P*
Between subjects	63			
Anxiety	1	1044·10	6·51	< 0·025*
Subjects within group	62	160·38		
Within subjects	255			
Visit	1	33·78		N.S.
Visit × Anxiety	1	71·19		N.S.
Visit × subjects within group	83	53·47		
Time	1	142·50	4·68	< 0·025*
Time × Anxiety	1	164·16	5·39	< 0·025*
Time × subjects within group	83	30·47		
Visit × Time	1	46·41		N.S.
Visit × Time × Anxiety	1	16·50		N.S.
Visit × Time × subjects within groups	83	32·95		

* One-tail.

No main effect or interaction involving 'Visit' is significant, supporting the argument that long-term habituation was not to be expected. The other expectations are reflected in the interaction term of Time × Anxiety, for both attention seeking and proximity seeking. As was expected the frequency of proximity seeking increased with time under increased level of anxiety, and decreased (only very slightly) under low-anxiety conditions. The interaction is significant ($P < 0·025$). Unlike proximity

seeking, attention seeking decreased with time (habituation) whatever the anxiety condition.

It is to be noted from Table 4, however, that the habituation was at a higher rate (steeper slope) under low anxiety than under high anxiety ($P < 0{\cdot}025$). This interaction was not expected. One could argue that anxiety level does after all have some effect on attention seeking since it inhibits to a certain degree its habituation. An alternative argument is that, by comparison with the low-anxiety condition, high anxiety had more 'novel' elements coming in at regular intervals, provoking comments which were scored as attention seeking. These interpretations are ad hoc and further research is needed to determine exactly the effect of anxiety on attention seeking. By comparison with the marked main effect of anxiety or proximity seeking ($P < 0{\cdot}025$), however, there was no such effect on attention seeking (Table 5).

Table 4. *Changes with time in the mean frequency of attention seeking and proximity seeking under two anxiety levels*

	Attention seeking				Proximity seeking			
	Time-1		Time-2		Time-1		Time-2	
	Mean	n	Mean	n	Mean	n	Mean	n
Low anxiety	12·19	64	8·89	64	6·62	64	6·52	64
High anxiety	11·56	64	10·73	64	9·06	64	12·16	64
	$P < 0{\cdot}025$				$P < 0{\cdot}025$			

Table 5. *The mean frequency of attention seeking and proximity seeking under two anxiety conditions*

	Attention seeking		Proximity seeking	
	Mean	n	Mean	n
Low anxiety	10·52	128	6·57	128
High anxiety	11·15	128	10·61	128

When habituation is considered regardless of anxiety condition (main effect of time), attention seeking proves to decline significantly with time, as was expected. Proximity seeking, however, increased significantly; this effect was not expected since the increase under high anxiety was expected to be balanced by the decrease under low anxiety; but this expectation was not confirmed, in spite of the fact that the expected interaction was significant and the order of the means was as expected. This result appears to be due to the fact that even under low anxiety proximity seeking showed only very slight habituation (see Table 4).

It could be that while reaction to novelty, as expressed in attention seeking, habituates quite quickly within a $\frac{1}{2}$ hr session, the reaction to the fear-provoking elements, as expressed by proximity seeking, takes much longer to habituate. Thus one would need a longer period than $\frac{1}{2}$ hr to habituate proximity-seeking behaviour effectively. Another possibility is that 'proximity seeking' is a personality characteristic which has a stable minimal level, and thus will never show habituation,

though it can be *increased* by an anxiety-provoking situation. This interpretation cannot be accepted, however, until it is convincingly demonstrated that proximity seeking does not habituate.

Discussion

The results confirm the basic expectations concerning habituation of dependency responses to an unfamiliar situation and the effects of an increased anxiety level on this habituation.

From the moment the child was approached by the experimenter she was confronted with an unfamiliar situation. She was told that she was going to play in a 'little nursery school' all of her own. The toys in the experimental room generated the familiar nursery-school atmosphere. These familiar elements, however, were in conflict with some incongruent elements in the situation, such as having to walk through a corridor of an office building or through an empty lecture hall en route to the experimental room, or having an office desk and a stationery cupboard behind the curtain of the 'little nursery school.' This incongruency was expected to generate some anxiety in the child, while the novel (but congruous) elements in the situation were expected to generate curiosity. It was also expected that both the reactions to incongruency and the reactions to novelty would habituate with continuous exposure to the situation. It was suggested furthermore that one type of dependency behaviour (attention seeking) would be affected by the novel elements in the situation, generating curiosity, while another type of dependency behaviour (proximity seeking) would be affected by the incongruous elements in the situation generating anxiety. It was finally suggested that a gradual introduction of anxiety-provoking elements into the situation would interfere with the habituation of proximity seeking (to the incongruous elements) so as gradually to increase its frequency. These additional anxiety elements were not expected to increase the frequency of attention seeking or interfere with its habituation.

As was expected, the results indicate that proximity seeking tends somewhat to habituate when no anxiety is added to the situation, but that its frequency increases with time when there is an increase in anxiety level. Attention seeking habituates under both conditions of anxiety. It habituates faster, however, when no additional anxiety element is introduced.

The investigation was supported by the Louis S. Haas Fund of Stanford University, California, U.S.A., and was carried out in partial fulfilment of the requirements for the Ph.D. degree of Stanford University. The author wishes to express her thanks to Robert R. Sears and Eleanor E. Maccoby of the Department of Psychology, Stanford University, for their patience, encouragement and constructive criticism. The author also appreciates the assistance and advice given by C. B. Hindley of the Centre for the Study of Human Development, University of London Institute of Education, and by A. R. Jonckheere of the Department of Psychology, University College London.

References

Arsenian, J. M. (1943). Young children in an insecure situation. *J. abnorm. soc. Psychol.* **38**, 225–49.

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

Fiske, D. W. & Maddi, S. R. (ed.) (1961). *Functions of Varied Experience*. Homewood, Ill.: Dorsey Press.

GLANZER, M. (1958). Curiosity, exploratory drive, and stimulus satiation. *Psychol. Bull.* **55**, 302–15.

HARLOW, H. F. (1953). The nature of love. *Am. Psychol.* **13**, 673–85.

HEBB, D. O. (1946). On the nature of fear. *Psychol. Rev.* **53**, 259–302.

HEBB, D. O. (1949). *The Organization of Behavior.* New York: Wiley.

HINDE, R. A. (1961). Changes of responsiveness consequent upon performance. In *Determinants of Infant Behavior*, ed. B. M. Foss, pp. 203–5. London: Methuen.

HUNT, J. McV. (1960). Experience and the development of motivation; some reinterpretations. *Child Dev.* **31**, 489–504.

ROSENTHAL, M. K. (1967). The generalization of dependency from mother to stranger. *J. Child Psychol. Psychiat.* (In the Press.)

SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., WAITE, R. R. & RUEBUSH, B. K. (1960). *Anxiety in Elementary School Children.* New York: Wiley.

WALTERS, R. H. & RAY, E. (1960). Anxiety, social isolation and reinforcer effectiveness. *J. Pers.* **28**, 358–67.

(*Manuscript received* 22 *March* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 365–373
Printed in Great Britain

TEMPORAL EFFECTS ON RESPONSE DECREMENT AND STIMULUS SATIATION IN EXPLORATION

By CORINNE HUTT
Park Hospital for Children, Oxford

Some workers have shown the temporal course of exploration to be decremented with time, while others have shown it to be maintained at a constant level from day to day. It has been argued that different behaviours are involved in the two cases and that the latter effect is a function of the experimental conditions. Experiments with nursery-school children substantiated this argument. A further experiment sought to establish the most appropriate theoretical construct to account for the decrement in exploration by manipulation of the temporal variable. It was found that two decremental processes were manifest: one due to response inhibition and the other to stimulus satiation. The former was seen to be primarily affected by the complexity of the novel stimulus rather than by temporal factors; the latter by both.

Studies concerning the temporal course of exploration have shown two contrary trends: on the one hand exploration has been shown to decrease with continued or repeated exposure, and on the other, to maintain a steady level over time. This discrepancy, though seldom made explicit, underlies an important difference in theoretical emphasis. Studies reporting the former trend tend to use a drive-reduction model to explain the decrement; those demonstrating the latter trend take it as being indicative of the intrinsic potency of exploration.

For example, Berlyne (1950) postulates that 'as a curiosity-arousing stimulus continues to affect an organism's receptors, curiosity will diminish'; moreover, his theory predicts a rapid but temporary decline in exploration within a single trial, and a more slowly developing permanent decline over repeated exposures, owing to inhibitory processes analogous to reactive and conditioned inhibition respectively. Many experiments with animals and adult human beings have verified these predictions.

On the contrary, Butler & Harlow (1954), Butler & Alexander (1955), Harlow, Blazek & McClearn (1956) have found that visual and manipulatory exploration in the monkey persisted over a long period of time and showed a non-decremental pattern with repeated trials. Similar results have been reported for rats (Kish & Antonitis, 1956), and for monkeys in other situations (Moon & Lodahl, 1956; Symmes, 1959).

The experimental conditions of the two groups of studies, however, were quite different: in the first, responses were elicited by novel or complex stimulations; in the second, some form of stimulation (not necessarily novel or complex) was contingent upon some response of the animal when in an unstimulating environment. Fowler (1965) has neatly summarized this distinction: in the one case the emphasis is on the animal making a consummatory response *to* a stimulus change, and in the other on making an instrumental response *for* a stimulus change. Myers & Miller (1954) have moreover questioned Harlow's interpretation of the same stimuli simultaneously eliciting the exploratory drive and serving as reinforcement: 'we suggest that drives produced by homogeneous or monotonous stimulation, enforced inaction, etc., may be reduced by sensory variety, freedom of action, etc. . . .'

Since therefore the non-decremental pattern of exploratory responsiveness seems to be largely a function of the unstimulating properties of the experimental situation, the provision of a realistically stimulating but familiar environment should result in a decline in exploration over time, even though specific stimulus changes are made contingent upon the exploratory responses. The first study was a test of this prediction. Young children were introduced individually into a familiar playroom in which a novel object was present, together with other familiar toys. The amount and nature of the child's responsiveness towards this object over repeated exposures to it formed the primary data of the study.

EXPERIMENT I

Method

Subjects. The subjects were eight nursery school children between the ages of 3 and 5 yr.

Apparatus. The novel object consisted of a box (12 in. × 6 in. × 6 in.) on four legs, a lever on top being movable in four directions; these movements were registered by four post-office counters which could be made visible or not. The design of the object also allowed the use of auditory incentive—a buzzer and a bell respectively which were contingent upon two specific manipulatory movements. Thus four incentive conditions differing in complexity were available: (i) no sound or vision—bell and buzzer switched off and counters covered; (ii) vision only—counters visible, noises off; (iii) sound only—bell and buzzer on, counters covered; (iv) sound and vision—bell and buzzer on, counters visible. In this experiment only conditions (i) and (ii) were used. Five other familiar and conventional commercial toys were also available to the child.

Procedure. Each child was given six exposures to the novel object. The six experimental sessions were preceded by two pre-exposure sessions, the latter serving to additionally familiarize the child with the room and toys. All sessions lasted 10 min, the pre-exposure sessions being procedurally identical with the experimental sessions except for the presence of the novel stimulus. The counter readings were taken before and after each experimental session. The time interval between consecutive sessions was 48 hr, except between sessions 3 and 4 when it was 72 hr owing to the intervention of the weekend. Behavioural data were recorded on 8 mm cine film or on check-lists (Hutt, 1966).

Results

Under the simplest incentive condition (i) there was a negatively accelerated decrement in responsiveness upon repeated exposures (Fig. 1). There was, however, a recovery of exploration in session 4, after the longer time interval. The counter readings were used as a measure of manipulatory exploration.

Under incentive condition (ii) a greater degree of stimulus change was effected by the lever manipulations; nevertheless, the prediction was that here too response decrement could be demonstrated. Furthermore, to test whether the recovery or disinhibition was in fact due to the greater time available for the dissipation of inhibitory activity, as opposed to some other factor operating spontaneously at the mid-point in the course of exploration, the weekend was interposed between sessions 4 and 5. Once again a decrement in exploration was observed with partial recovery occurring after the longer inter-session interval (Fig. 2). The McNemar test for the significances of changes on the combined data from (i) and (ii) (session before weekend × session after weekend) shows the recovery to be significant ($\chi^2 = 4{\cdot}16$,

D.F. 1, $P < 0{\cdot}05$, after correction for continuity). When the weekends were equally distributed over the six sessions for all the subjects, this response decrement was seen to be an exponential one under these two conditions (Hutt, 1966).

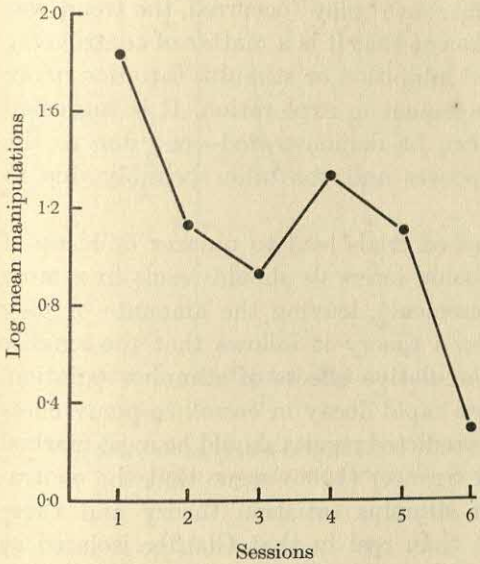


Fig. 1

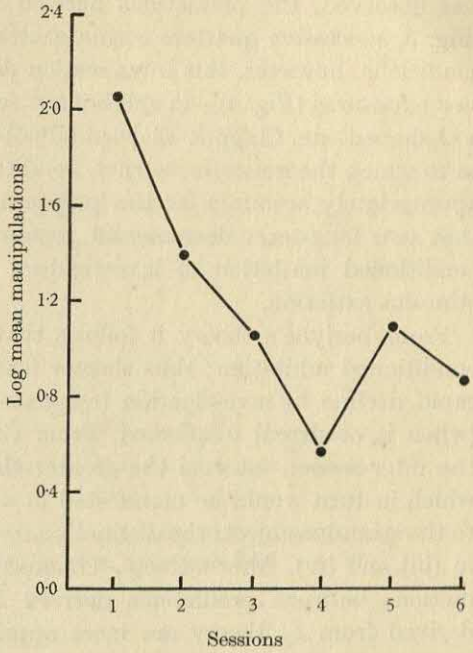


Fig. 2

Fig. 1. Temporal course of manipulatory exploration of a novel object when weekend (72 hr) intervened between sessions 3 and 4, other sessions being 48 hr apart ($n = 4$).

Fig. 2. Temporal course of manipulatory exploration when weekend intervened between sessions 4 and 5; visual feed-back of lever-operated counters was available ($n = 4$).

Experiment II

It had also been shown in the earlier study (Hutt, 1966) that two morphologically different categories of behaviour towards the novel object could be distinguished: investigation and 'other responses' involving the object. This latter category consisted of activities which would be termed 'play' by most observers. These two behaviours were seen to have different determinants and temporal courses, and it was suggested that they were also differently motivated. Play activities only appeared when auditory feed-back was available, i.e. under conditions (iii) and (iv).

Investigatory responses showed an overall decrement over the six sessions under both conditions (iii) and (iv), whereas play activities appeared to be a quadratic function of time, i.e. they increased progressively from session to session but began to show a decrease after the fifth exposure.

The different temporal characteristics of these two classes of behaviour were also reflected in their within-session trends. Berlyne's (1950) theory, with its emphasis on the response, predicts relatively rapid intra-session decrement in exploration

with a slower inter-session decrement. Glanzer's (1953) theory, emphasizing the stimulus aspects, predicts a linear decrease with time within sessions but not much inter-session decrement since stimulus satiation decreases exponentially with time in the absence of the stimulus. Under the no-sound conditions, where only investigation was observed, the predictions derived from Berlyne's theory are better fulfilled (Fig. 3, successive quarters within each session being considered). Under the sound conditions, however, this intra-session decrement was observed only upon the first two exposures (Fig. 4)—in subsequent sessions, once 'play' occurred, the trend was a U-shaped one. Clapp & Eichorn (1965) comment that it is a matter of controversy as to which theoretical construct, conditioned inhibition or stimulus satiation, more appropriately accounts for the long-term decrement in exploration. It is suggested that two long-term decremental processes can be demonstrated—one due to the conditioned inhibition of investigatory responses and the other possibly due to stimulus satiation.

From Berlyne's theory it follows that massed trials lead to quicker build-up of conditioned inhibition; thus shorter inter-session intervals should result in a more rapid decline in investigation (response decrement), leaving the amounts of play (when it occurred) unaffected. From Glanzer's theory it follows that the smaller the inter-session interval the greater the cumulative effects of stimulus satiation, which in turn would be manifested in a more rapid decay in *overall* responsiveness to the stimulus object; the distinction in the predicted results should be most marked in (iii) and (iv). Alternatively, Thompson & Spencer (1966) argue that the contradictions between predictions derived from stimulus satiation theory and those derived from I_R theory are more apparent than real in that Glanzer isolated as a response only one element, viz. the turn, from a whole response sequence. They conclude that 'insofar as behavioral responses are concerned, *the formal properties of stimulus satiation and response habituation are the same*' (italics mine). Thus, one would expect to find a unitary decremental process, and in this case the choice of which theoretical construct could best account for it would be largely one of personal preference.

Method

In order to test between these predictions two other temporal schedules were used in addition to 48 hr (alternate days): 24 hr (daily) and 12 hr (twice daily) intervals between successive exposures.

In all other respects the procedure was the same as for Expt I.

Procedure. The most simple (i) and the most complex (iv) incentive conditions were used.
Subjects. The subjects were twenty-four nursery school children between the ages of 3 and 5 yr.

Results

Under condition (i), no sound or vision, response decrement on the three temporal schedules was similar (Fig. 5). The data were transformed using a Naperian logarithmic scale. Analysis of variance for trend (Alexander, 1946) showed no significant difference between alternate-day, daily, and twice-daily exposures (Table 1A). The temporal effects were more evident, however, under condition (iv), sound and vision (Fig. 6); overall responsiveness toward the object reached a peak sooner and decreased

to a lower level with twice daily exposures than with daily or alternate-day sessions. When, however, the amounts of investigatory response are considered under the three temporal conditions, the decrements appear to be essentially similar (Fig. 7, Table

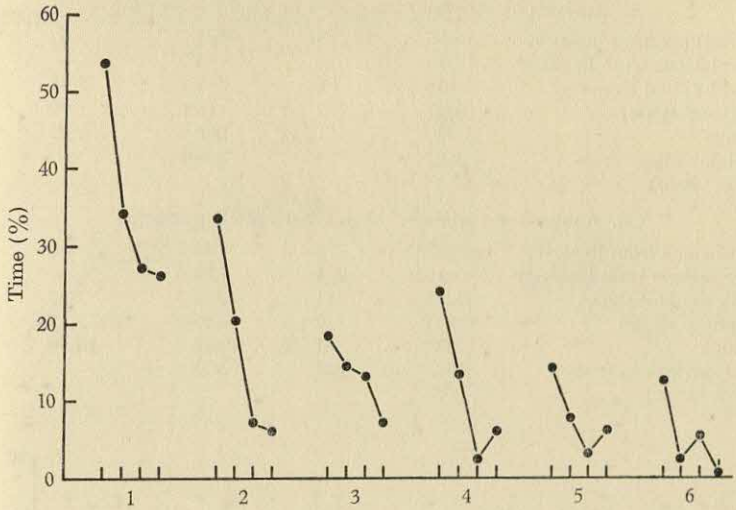


Fig. 3. Inter- and intra-session decrements in exploration when no auditory feed-back was available, successive quarters ($2\frac{1}{2}$ min) of each session being considered ($n = 13$).

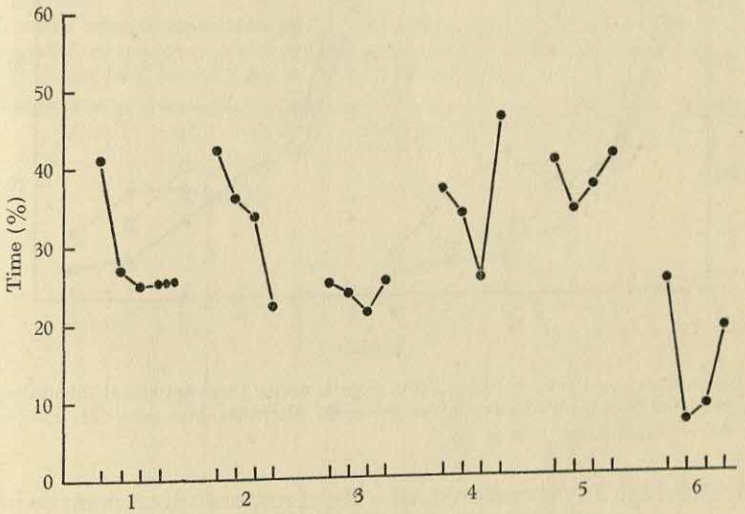


Fig. 4. Inter- and intra-session decrements in exploration when response-contingent auditory feed-back was available ($n = 12$).

1B). Since both 'group deviations from estimation' and 'overall deviations from linearity' are smaller than the error term (individual deviations from linearity) it can be concluded that response decrement shows no systematic deviation from linearity. The component 'between group slopes' is not significantly greater than error, thus indicating that the rates of decrement in the three groups are not

TABLE 1

Source	Sums of squares	D.F.	Mean squares	F	P
A. Analysis of variance: 'no sound or vision' condition					
Group deviations from linearity	868	8	108·5	—	—
Overall deviations from linearity	349	4	87·2	—	—
Between individual slopes	2408	7	344·0	2·48	< 0·05
Between group slopes	21	2	10·5	< 1·00	N.S.
Overall slope	960	1	960·0	6·92	< 0·05
Individual deviations from estimation (error)	3881	28	138·6	—	—
B. Analysis of variance: 'sound and vision' condition					
Group deviations from linearity	1083	8	135·3	—	—
Overall deviations from linearity	618	4	154·5	—	—
Between individual slopes	3295	11	299·5	1·67	N.S.
Between group slopes	1047	2	523·0	2·93	N.S.
Overall slope	2332	1	2332·0	13·00	< 0·01
Individual deviations from estimation (error)	7858	44	178·6	—	—

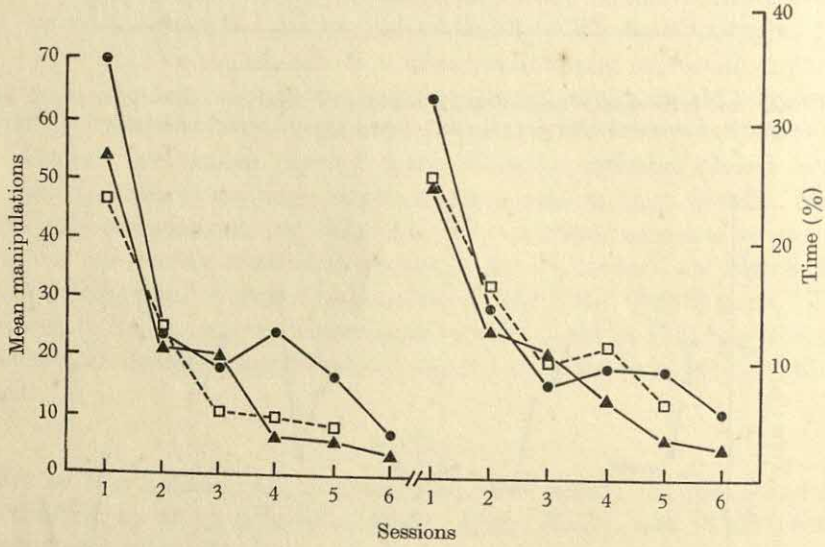


Fig. 5. Manipulatory and total exploration of object, under three temporal schedules when no auditory or visual feed-back was available: ●——●, alternate days ($n = 3$); □——□, daily ($n = 3$); ▲——▲, twice daily ($n = 4$).

significantly different. In other words, the rate of response decrement or investigation was not appreciably affected by massed as opposed to distributed trials. On the other hand, there appeared to be marked differences in the amounts of play activities engaged in under the three temporal conditions (Fig. 8).

Analysis of variance on the proportions of time spent in play showed the curves for the daily and twice-daily conditions to be significant quadratic functions of the number of exposures ($F = 990{\cdot}5/108{\cdot}4$, D.F. 1, 18, $P < 0{\cdot}01$, and $F = 349{\cdot}9/70{\cdot}4$, D.F. 1, 18, $P < 0{\cdot}05$ respectively). The quadratic component of the alternate-day curve fell short of significance. The null hypothesis, that the maxima under the

three temporal conditions were not significantly different from each other, was tested by the Kruskal–Wallis one-way analysis of variance. This hypothesis was rejected at the 5% level ($H = 6{\cdot}01$; D.F. 2). It can be concluded therefore that the amounts of play under the three temporal conditions, as reflected by their maxima, are significantly different from each other.

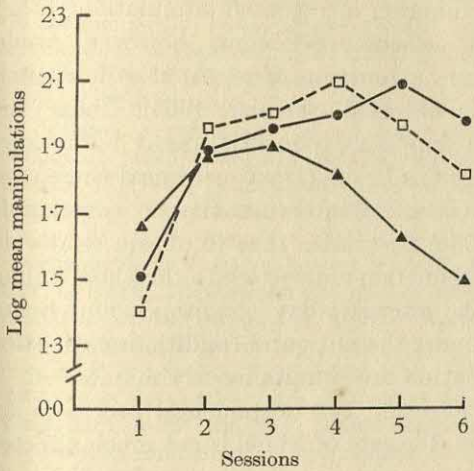


Fig. 6

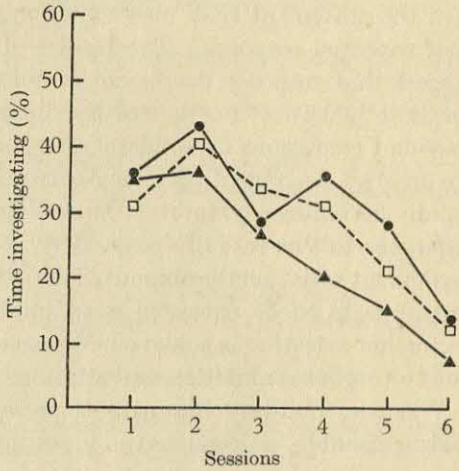


Fig. 7

Fig. 6. Overall responsiveness towards novel object under three temporal schedules when response-contingent auditory and visual feed-back were available: ●——●, alternate days ($n = 6$); □——□, daily ($n = 4$); ▲—▲, twice daily ($n = 4$).

Fig. 7. Decrements in investigatory responses under three temporal schedules: ●—●, alternate days ($n = 6$); □—□, daily ($n = 4$); ▲—▲, twice daily ($n = 4$).

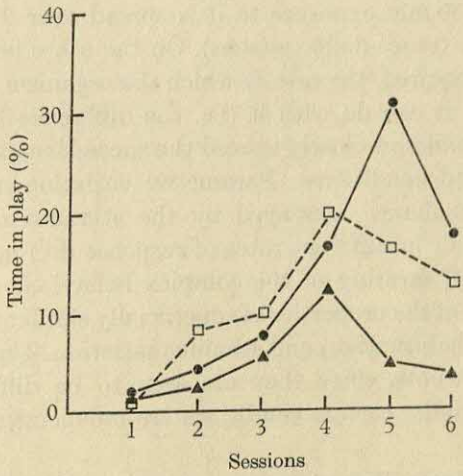


Fig. 8. Amounts of 'play' involving the object displayed under three temporal schedules, with sound and vision: ●—●, alternate days ($n = 6$); □—□, daily ($n = 4$, except for sixth session when data for only two subjects); ▲—▲, twice daily ($n = 4$).

DISCUSSION

Traditionally, the theoretical debate on extinction centres on the issue of whether reactive inhibition, i.e. an inhibitory process operating at the effector end, or stimulus satiation, i.e. an inhibitory process operating at the receptor end of the S-R chain, more satisfactorily accounts for the phenomena observed (Glanzer, 1953; Broadbent, 1958; Clapp & Eichorn, 1965). Broadbent concludes a discussion of vigilance tasks with the statement that 'decrement was a function of repeated stimulation rather than repeated response'. The results of the present experiments, however, would suggest that response decrement is not purely a function of repeated stimulation, nor is it 'solely a function of the inter-trial interval' (Fowler, 1965). Under the no-sound conditions investigatory responses show an exponential decay with time, the greatest decrement occurring between sessions 1 and 2, responsiveness thereafter rapidly becoming asymptotic. Decreasing the intertrial interval makes no appreciable difference to this rate of decay. Berlyne (1955) postulates that to obtain relatively permanent satiation the amount of curiosity reduction must exceed a threshold value; this appears to be exceeded even under the alternate-day conditions, and hence no further satiation is achieved. Moreover, under the no-sound conditions extinction due to response inhibition and stimulus satiation are simultaneously manifested.

When an auditory incentive is present, however, two decremental processes are distinguishable: a decrement in investigation, the rate of which is not much affected by massed trials; and a decrease in overall responsiveness due to stimulus satiation. The effects of the latter are mainly reflected in the amounts of play engaged in with the object.

These results suggest that the rate at which response decrement occurs is more a function of stimulus complexity and novelty than of temporal factors. In other words, from a given set of stimulus parameters there seems to be a certain amount of information to be gained and the rate at which the organism does this is essentially the same whether the 60 min exposure to it is spread over 2 weeks (alternate-day sessions) or over 3 days (twice-daily sessions). On the other hand, once most of this information has been acquired, the rate at which the organism 'tires' of the stimulus depends on how much it can do with it (i.e. the difference between the no-sound and sound conditions) and how closely spaced the successive exposures are (temporal effects within the sound conditions). Parametric variation of the information, or conversely of the redundancy, conveyed by the stimulus object should, on the basis of these data, result in different rates of response decrement.

A close morphological scrutiny of the complex behaviour of young humans has enabled the elucidation of the properties of superficially similar theoretical constructs: response inhibition (or habituation) and stimulus satiation. These cannot be regarded as interchangeable concepts, since they are seen to be differentially affected by temporal factors, and under certain conditions are independent of each other.

I am grateful to John Bevan, Unit of Biometry, to Dr J. A. Gray, Institute of Experimental Psychology, and to S. J. Hutt, St Catherine's College, for their helpful criticism and statistical advice. This study was made possible by the generous support of the Nuffield Foundation and the kind co-operation of Miss J. Lawrence and her staff at the Slade Nursery School, Oxford.

References

Alexander, H. W. (1946). A general test for trend. *Psychol. Bull.* **43**, 533–57.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Br. J. Psychol.* **41**, 68–80.

Berlyne, D. E. (1955). The arousal and satiation of perceptual curiosity in the rat. *J. comp. physiol. Psychol.* **48**, 238–46.

Broadbent, D. E. (1958). *Perception and Communication*. Oxford: Pergamon Press.

Butler, R. A. & Alexander, H. M. (1955). Daily patterns of visual exploratory behaviour in the monkey. *J. comp. physiol. Psychol.* **48**, 247–9.

Butler, R. A. & Harlow, H. F. (1954). Persistence of visual exploration in monkeys. *J. comp. physiol. Psychol.* **47**, 258–63.

Clapp, W. F. & Eichorn, D. H. (1965). Some determinants of perceptual investigatory response in children. *J. exp. Child Psychol.* **2**, 371–88.

Fowler, H. (1965). *Curiosity and Exploratory Behavior*. New York: Macmillan.

Glanzer, M. (1953). Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.* **60**, 257–68.

Harlow, H. F., Blazek, N. C. & McClearn, G. E. (1956). Manipulatory motivation in infant rhesus monkeys. *J. comp. physiol. Psychol.* **49**, 444–8.

Hutt, C. (1966). Exploration and play in children. *Symp. Zool. Soc. Lond.* **18**, 61–81.

Kish, G. B. & Antonitis, J. J. (1956). Unconditioned operant behaviour in two homozygous strains of mice. *J. genet. Psychol.* **88**, 121–9.

Moon, L. E. & Lodahl, T. M. (1956). The reinforcing effect of changes in illumination on lever-pressing in the monkey. *Am. J. Psychol.* **64**, 288–90.

Myers, A. K. & Miller, N. E. (1954). Failure to find a learned drive based on hunger: evidence for learning motivated by 'exploration'. *J. comp. physiol. Psychol.* **47**, 428–36.

Symmes, D. (1959). Anxiety reduction and novelty as goals of visual exploration by monkeys. *J. genet. Psychol.* **94**, 181–98.

Thompson, R. F. & Spencer, W. A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol. Rev.* **73**, 16–43.

(*Manuscript received* 26 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 375–386
Printed in Great Britain

COMPONENTS OF CURIOSITY: BERLYNE RECONSIDERED

By ELLEN GREENBERGER,* JANET WOLDMAN† and SARAH WILLIAMS YOURSHAW

Wellesley College, Mass., U.S.A.

Berlyne has found that collative stimuli—stimuli with the attributes of complexity, incongruity, or ambiguity, among others—attract longer visual exploration than their less perplexing counterparts. He has postulated a curiosity motive to account for the finding; this paper suggests an alternative explanation of some of his results. The first of two experiments evaluates whether viewing time allotted to collative stimuli is increased by prior curiosity-arousal. The second evaluates whether a 'set to remember' also increases viewing of collative stimuli. Results indicate (*a*) that both experimental conditions elevate attention to collative stimuli, and to approximately the same extent; but (*b*) that curiosity-arousal particularly elevates attention to incongruous stimuli, whereas, a 'set to remember' increases attention to complex stimuli. Control subjects given instructions patterned after Berlyne's seem to exhibit a self-induced remembering set. This emphasizes the need to separate the effects due to curiosity from those due to a remembering orientation.

That it takes something special to arouse a person's curiosity is understood by any writer who has ever agonized over the first line of a paper. The systematic study of curiosity, however, has gained momentum only during the past decade. Research efforts in this period have investigated such issues as why the rat varies his route through a maze (Dember, 1956, 1961); why organisms look at one aspect of the environment in preference to another (Berlyne, 1950, 1954*b*, 1957*b*, 1958*a*, *b*; Cantor, Cantor & Ditrichs, 1963; Fantz, 1961); which objects organisms play with (Mendel, 1965; Welker, 1956*a*, *b*); and which things people want to know more about (Berlyne, 1954*a*, *b*, *c*). The diversity of behaviours included in the literature on curiosity is immediately apparent: alternation behaviour, visual exploration, manipulation, play, and the quest for knowledge. Whether these and related behaviours reflect a common underlying process, and whether this process should be called a motive, are questions that now are beginning to receive attention (Berlyne, 1963*a*; Fiske & Maddi, 1961; Maddi, Propst & Feldinger, 1965; Rapaport, 1960).

Berlyne's theory of curiosity motivation

Berlyne maintains that certain of the behaviours just mentioned reflect a boredom drive, while others are the outgrowth of a curiosity drive. The concept of a boredom drive is similar to the 'optimal level of excitation' proposed by Hebb & Thompson (1954) and developed by Leuba (1955), and it is also similar in certain respects to the 'need for variety' postulated by Fiske & Maddi (1961) and by Maddi, Propst & Feldinger (1965). Berlyne assumes a boredom drive to underly exploratory activities for which 'any stimulus that is sufficiently interesting or entertaining' is reinforcing (Berlyne, 1963). That is, the organism does not seek contact with any particular object; he seeks a certain broad type of stimulation which may be encountered in a variety of situations. The spirit of such 'diversive exploration' is aptly phrased by

* Now at the Center for the Study of Social Organization of Schools and the Learning Process and in the Department of Social Relations, The Johns Hopkins University.

† Now attending the University of Kansas.

the bored child who asks, 'What shall I do now, mommy?' A curiosity drive is invoked in cases of specific exploration. Specific exploration has the function of obtaining stimulation from a definite source. No other will do. It is the kind of behaviour we engage in when looking for the star in the sapphire, or looking for the needle in the haystack.

Berlyne further distinguishes between extrinsic and intrinsic curiosity (Berlyne, 1960, 1963*a*). In the former case, exploration is instrumental to the attainment of some other goal; it paves the way for a subsequent response with its own source of reinforcement. Intrinsic curiosity involves exploration 'for its own sake' (perhaps what the layman calls 'idle' curiosity), and appears to be rewarding in its own right (Butler, 1953; Harlow, Harlow & Meyer, 1950).

Berlyne sets curiosity in the framework of Hullian drive theory. He argues that there is in both man and lower animals a drive that is aroused by conflict. Conflict, in turn, creates tension, and exploratory behaviour is an attempt to dispel tension. The resolution of conflict is tension-reducing (rewarding). One group of stimulus-characteristics, called *collative* properties, is especially effective in provoking conflict and subsequent exploratory behaviour. These stimulus-properties include novelty, surprise, change, ambiguity, incongruity, complexity, blurredness, and power to induce uncertainty (Berlyne, 1960, 1963). The term 'collative' indicates that these stimuli induce the viewer to collate or compare information from several sources. For example, a novel stimulus is novel to us only insofar as we compare it with other stimuli present at that time or with stimuli we have encountered in the past.

Conflict consists in the simultaneous arousal of incompatible responses. A vivid, though inelegant, example from the comic strips illustrates how conflict is triggered by incongruity. As the figure of the hero is seen in the sky, curious onlookers cry out, 'It's a *bird*! It's a *plane*!...' Additional responses associated with bird and plane are instigated by stimulus generalization, and these responses, like 'bird' and 'plane', also prove incompatible to some degree. Exploration results in increased familiarity with the stimulus object. As a consequence, some response gains preponderance over the others: 'It's *Superman*!' Conflict is reduced, and exploration (curiosity) gradually abates.

Most of Berlyne's experimental work has dealt with perceptual curiosity (e.g. Berlyne, 1954*a*, 1957*b*, 1958*a*, *b*, *c*, 1963*b*) which he defines as a drive aroused by a conflict-producing visual stimulus and reduced by continued perceptual contact with that stimulus. Differences in the duration of perceptual contact (hereafter called *viewing time*) should reflect both the degree to which the stimulus creates conflict and the strength of the organism's exploratory response. (The former issue is subsumed under the Drive variable in Hull's model; the latter under Habit Strength.) In a number of experiments using viewing time as the dependent variable, Berlyne has demonstrated the power of collative stimuli to attract visual exploration.

The following experiments illustrate Berlyne's procedures and findings, and are the starting point for questions that are raised in the remaining sections of this paper.

In an experiment with adults, Berlyne (1958*a*) exposed subjects to pairs of figures, each appearing for 10 sec. Four sets of figures contained pairs whose members varied in complexity (i.e. novelty, in terms of information theory concepts). Complexity was defined in terms of irregularity of shape, amount of material, and heterogeneity

of elements (see Berlyne, 1960, for more complete definitions of collative stimulus-properties). The figures, with the exception of one, were of abstract rather than concrete forms. In the remaining two sets of figures, an incongruous figure was paired with a congruous one. Incongruity involved stimulus-patterns in which one part conflicted with expectations aroused by the other: thus, an elephant's head on a body with lion's legs. All the stimuli were of concrete things, mainly animals. Berlyne found that for every pair of stimuli in the series, the more complex or incongruous figure was fixated for a significantly higher proportion of the 10-sec viewing period. A similar experiment (Berlyne, 1958*c*) using a 2-min exposure time yielded the same results.

In another experiment, Berlyne (1957*b*) investigated stimuli with the properties of incongruity, surprise, and complexity. The incongruous stimuli were pictures of animals and birds, as described before. The surprising stimuli consisted in a series of twelve abstract shapes, where the first six were red, the next five were green, and the twelfth was violet. Surprise was imputed to numbers 7 and 12. Complexity of two sorts was evaluated (see Berlyne, 1960). Amount of material, heterogeneity of elements, and irregularity of shape again were the dimensions of complexity involved. In this study, unlike the first, objects as well as abstractions were shown. The subject was told that the experiment was concerned solely with finding out how interesting certain pictures were, and that no questions about the figures would be asked at any time. The subject operated a tachistoscope, set for 0·14 sec exposures, and could press the key as many times as he wished to see each figure. Results indicated that stimuli with higher collative properties attracted significantly longer viewing time.

An alternative explanation of Berlyne's findings

These experiments raise some questions of interpretation and conceptualization, among them this one: Do Berlyne's subjects look longer at the collative stimuli because a *curiosity motive* has been aroused or because a *remembering set* has been evoked?

It is true that subjects were told that they would not be asked any questions about the figures, but subjects in psychological experiments often choose to disbelieve the experimenter; or at any rate, to make their own guesses regarding 'what the experiment is all about' (Orne, 1962). It seems plausible that a self-induced remembering set would draw attention to the same stimulus-properties that Berlyne believes engage the curiosity motive. The complex stimulus has more elements, is less regular, and is less easily labelled than its simpler mate. It is therefore harder to remember, and hence might be viewed longer. The ambiguous figure must be identified to be recalled, and thus might be viewed longer than its unambiguous mate. Finally, the incongruous stimulus requires identification of both the object and its incongruous features, resulting in longer viewing.

Rapaport (1960) also has questioned whether Berlyne's experiments reflect a curiosity motive. He claimed that Berlyne's results can be explained by a non-motivational theory of structure-building. This theory, based on Freud's uncompleted theory of consciousness, suggests that it is the development and maintenance of *structures* which underlies all the active, exploratory, or 'curious' modes of behaviour mentioned at the beginning of this paper. The key concepts are schema, attention, and structure-building.

Like Piaget, Rapaport (1960) conceptualized exploratory behaviour as the adaptation of an existing schema to a new object. When the schema has reached equilibrium, it has become a new structure. Exposure to new objects disturbs the schema's equilibrium again and initiates its differentiation into new schemata. 'Attention cathexis attracted by an excitation in a sufficient amount and for a sufficient duration gives rise to a structure (e.g. memory trace)' (Rapaport, 1960). Rapaport went on to make a number of specific proposals about the length of time different stimuli are explored. Two of them may be paraphrased as follows. (1) If a structure corresponding to past stimulus experience exists, external excitation will attract attention, but attention will soon become available again to other excitations. (2) However, if no structure exists and if the external stimulation is novel, structure-building requiring more intense and longer attention will take place. Rapaport's theory is pertinent to the previous discussion of implicit remembering sets triggered by the experimental task. In a sense, Rapaport maintains that processes conducive to remembering (i.e. structure-building) go on more or less automatically and continuously. The organism is 'set' by the nature of consciousness to attend to novel experience, and to be capable, therefore, of remembering. We will adhere to the one meaning of 'structure' explicitly cited—*memory trace*—in discussing how Rapaport's theory fits with Berlyne's findings.

Because of our familiarity with certain simple abstract figures, these figures may already be represented internally by a corresponding trace, and thus attract only limited attention. If not, structure-building will take longer in relation to *complex* figures than in relation to more regular, repetitive stimuli. The same lines of thought, concerning (*a*) the prior existence of structures and (*b*) the relative ease of structure-building, lead to the prediction of longer attention to *ambiguous* and *incongruous* stimuli than to their non-collative mates.

In summary, Berlyne's theory of curiosity motivation, Rapaport's non-motivational theory of structure-building, and the allied notion of subjects' self-induced remembering sets all predict similar behaviour with respect to collative stimuli. Three questions in need of clarification were investigated in the experiments reported in the next section: (1) Does attention to collative stimuli reflect curiosity, as conceptualized by Berlyne? (2) Does attention to collative stimuli reflect a set to remember, or the process of structure-building? (3) Do curiosity-arousal and the evocation of a remembering set have the same effect on attention to collative stimuli?

EXPERIMENTS I AND II

Experiment I

If Berlyne's and analogous procedures tap the curiosity motive, subjects previously exposed to a task that arouses curiosity should score higher than subjects given a non-arousing task. (An analogous argument was used by McClelland *et al.* (1953) in developing a measure of the achievement motive.) Put another way, if viewing time accorded to collative stimuli reflects curiosity motivation, length of viewing time should sensitively reflect the presence or absence of the motive or its variations in strength.

Method

Plan and materials

Curiosity was aroused in a manner consistent with Berlyne's (1960) definition of epistemic curiosity. The technique was based on that of Mittman & Terrell (1964), with modifications introduced in view of the difference in age of the subjects.

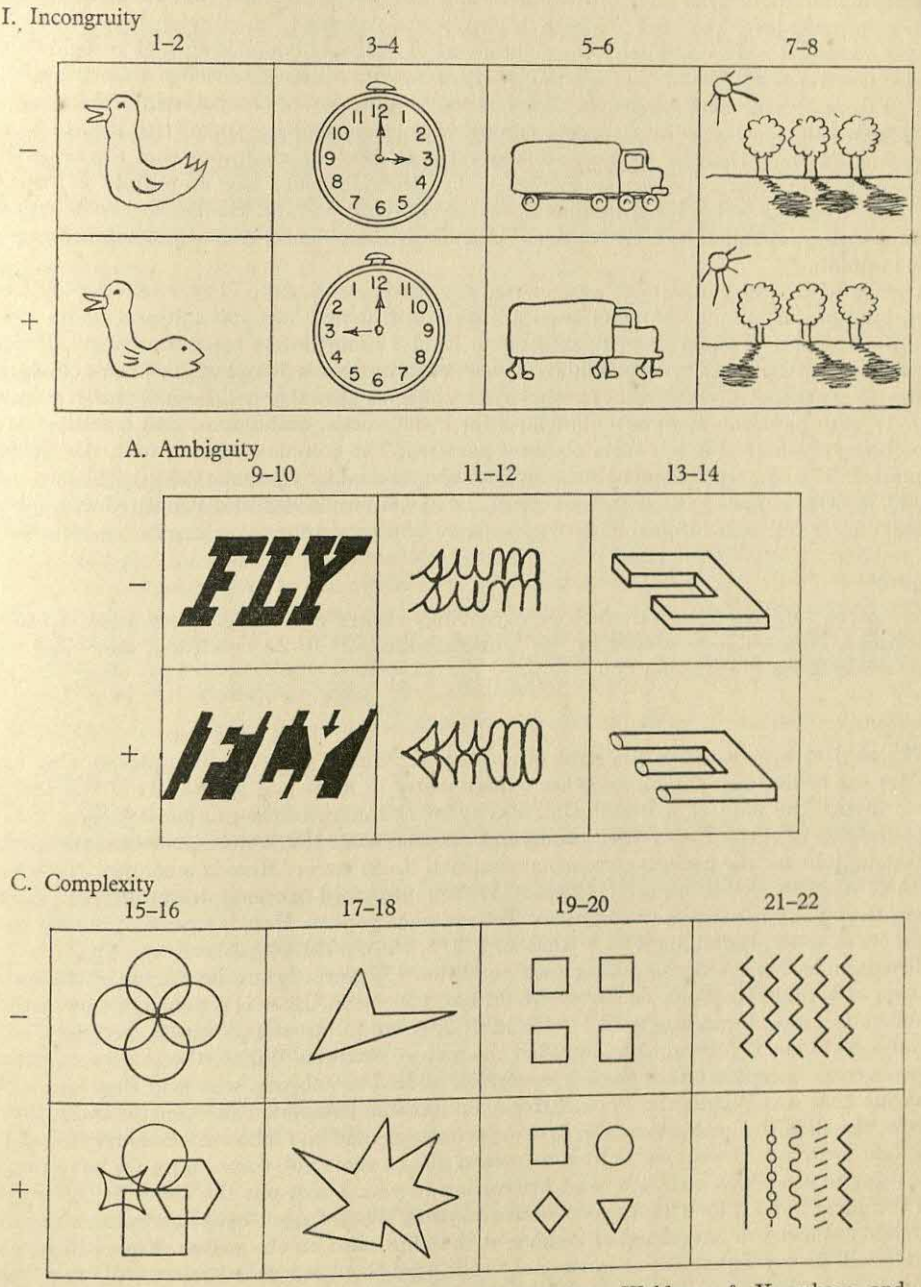


Fig. 1. Stimuli to measure curiosity devised by Greenberger, Woldman & Yourshaw, and based on Berlyne (1960). Plus signs indicate the member of each pair of stimuli with more collative properties.

Subjects were given a connect-the-numbers task which when completed constitutes a picture of Winnie the Pooh. The picture is constructed of odd numbers only, but some even numbers are present to help mask the identity of the picture. (Subjects in a pilot investigation all reported that they could not tell what the picture was by just looking at the numbers.) Subjects in the non-arousing task condition (T) were told what the picture was before being asked to connect the numbers. Subjects in the curiosity-arousal condition (C-A) were not told, and were asked to keep trying to guess what the picture was. Since the first group was told the identity of the picture immediately, the task exposed them to little conflict, as previously defined. With curiosity aroused only minimally, the reinforcement value of completing and recognizing the picture was relatively small. The second group, however, whose knowledge was delayed, had time to form several conflicting hypotheses about the identity of the pattern. Since this group had a higher degree of curiosity-arousal, completing and recognizing the picture should be more rewarding for them than for the non-aroused subjects. Berlyne maintains that the arousal and subsequent reduction of curiosity increases, by generalization, the amplitude of 'curious' responses to similar stimulus situations in future (Berlyne, 1954*b*). Hence, it is to be expected that the C-A group might be more 'curious' in the subsequent encounter with conflict-producing visual stimuli.

A series of twenty-two stimuli was devised to measure curiosity. They were selected from a larger number of stimuli that had been pre-tested on more than 100 subjects under neutral conditions similar to those given to subjects in Expt I immediately prior to viewing the slides. For each of the final twenty-two stimuli (eleven collative figures paired with eleven non-collative figures) a consistent preference in favour of the collative stimulus had been found. The stimuli (Fig. 1) were based on Berlyne's guidelines for incongruous, ambiguous, and complex figures, some being concrete things, others abstract patterns. The complex stimuli, with the exception of numbers 17 and 18, are exact or close replicas of ones used by Berlyne (1958*a*). They are shown in Fig. 1. The stimuli were drawn on 3¼ in. × 4 in. lantern slides and projected in a different random order for each subject. The two members of a pair were never shown consecutively.

Subjects

Seventeen subjects from introductory psychology classes were assigned randomly to the two conditions. Nine subjects served in the 'curiosity-arousal' (C-A) condition, and eight in the 'non-arousing task' condition (T).

Procedure

The subject was seated to the right of a table on which a projector was placed. One experimenter sat behind this table, the other experimenter in front and to the left of the projector. After giving the subject a pencil and a copy of the connect-the-numbers picture, the first experimenter read the proper instructions and watched while the subject connected the numbers.

Instructions for the curiosity-arousing condition (C-A) were: 'Here is a connect-the-numbers picture. What do you think it is?' (Pause) 'We are interested in seeing what different pictures it seems to you as you connect the numbers. Tell me what you see. Here is a pencil. Connect the odd numbers in order, beginning with 1 (that is, 1, 3, 5, etc.), until the picture is complete.'

Instructions for the non-arousing task condition (T) were: 'Here is a connect-the-numbers picture of Winnie the Pooh which we would like you to do. Here is a pencil. Connect the odd numbers in order, beginning with 1 (that is, 1, 3, 5, etc.) until the picture is complete.'

Subjects in the C-A condition identified the picture either as Winnie the Pooh or a bear. The experimenter accepted either answer as correct, adding to subjects who said they saw a 'bear' that the bear was Winnie the Pooh. After a subject had completed the picture and knew what it was, the slide test was given. The first experimenter read the following instructions and gave the light switch to the subject: 'Now we would like to show you some slides we have prepared. We want to see which ones are most interesting to you. I will put the slides in the projector. All you have to do is look at them at your own pace. When I say "okay", would you switch on the light? When you have finished looking at the slide, turn off the switch. Your turning off the switch will be my signal to put in the next slide, and then you wait again until I say "okay". This is not a speed test. We merely want to see which pictures you find most interesting, so you may take as long as you want.' These instructions were the same for both C-A and T groups. The second experimenter used a stopwatch to record the length of time between subjects turning the

projector on and off for each slide. (The timer attended to the clicking of the switch and did not see the stimulus.) The experimenter and her timing and recording activities could be seen by the subject. In order to control for experimenter error, the two experimenters changed jobs on the second day of experimenting.

Experiment II

As has been argued, it seemed possible that longer attention to collative stimuli might reflect (*a*) subjects' conscious or pre-conscious intent to memorize; and/or (*b*) the automatic, more or less continuous process of 'structure-building'. Expt II examined the effect of an experimentally induced remembering set, contrasted with a non-remembering, control set and explored the possibility that a conscious or automatic effort to remember may have been present in Berlyne's subjects. Indirectly, through the questions described below, the experiment touches on Rapaport's structure-building theory.

Method

Subjects

As in the first study, subjects were students in introductory psychology. Nine were assigned at random to the remembering condition (R); ten to the control condition (C).

Materials

The slide series shown in Fig. 1 again was used to measure viewing time.

In order to determine whether the control group was in fact neutral, or had a self-imposed set not unlike that induced in the experimental group, three questions were asked after the slides had been viewed. The first was an open-ended question permitting the subject to describe her conscious or preconscious set in her own words. The second was a forced-choice question, presumably tapping the same variable: items focused on curiosity and remembering sets of various kinds. The third question, also forced-choice, required the subject to report on her behaviour while looking at the stimuli: items focused on behaviours conducive to remembering and therefore, presumably, to structure-building. The first question was:

What did you think our purpose was in showing you these slides? Be specific.

The second question was:

What did you think our purpose was in showing you these slides? Check as many of the following as occurred to you. If none of these thoughts occurred to you, check the first item only.

1. I thought you were timing how long I looked at each slide to see which ones I found most interesting.
2. I thought you were going to ask me to recall some of them.
3. I thought you were going to ask me which ones I found most interesting.
4. I thought you were going to ask me which ones I found most aesthetic.
5. I thought you were going to ask me to draw some of them.
6. I thought it was a test of curiosity.
7. I thought it was a test of personality traits.
8. I thought it was a test of intelligence.
9. I thought it was a test of memory.

The third question was:

Please try to describe what you were 'doing' as you looked at the slides. Check as many of the following as apply.

1. I just looked at them until I lost interest.
2. I looked at them with the plan of being able to describe them later.
3. I tried to make a mental 'photograph' or visual image of the slides: i.e. more or less memorize them.

4. I labelled them as I went along: e.g. 'four squares'.
5. I tried to decide whether I had seen the slide before.
6. I compared the slide with others in the series, or with things I've seen in the past.

Procedure

Instructions for the remembering set condition (R) were: 'We have prepared a group of slides which we would like to show to you. After you have looked at them, you will have another task which will require that you remember some of them. I will put the slides in the projector. All you have to do is look at them at your own pace.' [The rest of the instructions were identical to the instructions for Expt I, except for the last line.] 'This is not a speed test. But we are going to ask you to recall some of the slides.'

Instructions for the control condition (C) were: 'We have prepared a group of slides which we would like to show to you. This experiment is intended to show how interesting certain pictures are. All you have to do is look at the slides at your own pace, and we will keep track of how long you look at each. I will put the slides in the projector.' [The rest of the instructions were identical to Expt I, except for the final two lines.] 'Remember, the experiment is concerned solely with showing how interesting certain pictures are. No questions about the pictures will be asked at any time.'

After viewing the slides, subjects in the remembering condition were told that no additional task was required after all; subjects in the control condition were given the questions to answer. The questions were given one at a time to prevent response-contamination.

Results

Overall preference for collative stimuli

Although other subjects had demonstrated longer attention to the collative member of each pair of stimuli, the question was evaluated again. Subjects from all four treatment groups were combined and, for each stimulus-pair, the number of times the collative stimulus was viewed longer was compared with number of preferences for the non-collative stimulus. (Ties were eliminated by discarding the few instances in which they occurred.) The more collative stimulus was viewed longer in every case. The binomial probabilities corresponding to the observed frequencies were as follows: for six slide pairs, $P < 0{\cdot}01$; for four pairs, $P < 0{\cdot}05 > 0{\cdot}01$; and for the one remaining pair, $P < 0{\cdot}1 > 0{\cdot}05$.

Effect of experimental conditions

For each pair of stimuli, the proportion of time each subject devoted to the collative member was found. Each subject's final score was the average proportion of time allotted to the eleven collative stimuli. Group results, expressed in percentages, are given in Table 1. The first row of figures indicates that the curiosity-arousal group looked longest at the collative stimuli, followed very closely by the remembering group, and finally (in a tie) by the task and the control groups. The scores of the task and control groups are remarkably similar in every respect, as Table 1 shows. A Mann–Whitney test confirmed this impression: the between groups difference is negligible ($z = 0{\cdot}31$; $P > 0{\cdot}3$). Because of the similarity in scores, and because T and C were both, in fact, 'control' groups, subjects from these two conditions were combined in order to increase the size of the group providing baseline data. To avoid confusion, the term *comparison group* will be used to refer to the combined task and control groups.

The questions listed at the end of the introduction can now be answered. Mann–Whitney tests were applied in each case.

1. Attention to collative stimuli does reflect curiosity. Subjects in the curiosity-arousal condition looked significantly longer at the collative stimuli than subjects in the comparison group ($z = 2{\cdot}93$; $P < 0{\cdot}001$).

2. Attention to collative stimuli does reflect a set to remember. Subjects in the remembering condition looked significantly longer at the collative stimuli than comparison subjects ($z = 2{\cdot}60$; $P < 0{\cdot}005$).

3. Curiosity-arousal and an experimentally induced set to remember have about the same overall effect on attention to collative stimuli; means for the entire series of stimuli were about the same. Differences between the two experimental groups, C-A and R, were non-significant ($z = 0{\cdot}84$; $P > 0{\cdot}4$). However, the two procedures do not affect the different types of stimuli equally, as is pointed out in the discussion, below.

Table 1. *Mean percentage of total viewing time allotted to the more collative member of each pair of stimuli under four conditions: curiosity-arousal (C-A) and non-arousing task (T) conditions in Expt I; and remembering (R) and control (C) conditions in Expt II*

	Experiment I		Experiment II	
	C-A	*T*	*R*	*C*
n	9	8	9	10
Mean for entire series	60	56	59	56
Incongruous stimuli	61	54	54	55
Complex stimuli	57	56	63	55
Ambiguous stimuli	62	60	60	59
Mean total viewing time (sec)*	102	73·1	237	75·6

* Subjects in the C-A condition spend more time looking at both collative and non-collative stimuli than subjects in the T condition. Subjects in the R condition likewise looked longer at each class of stimuli than their controls.

Analysis of questions

The questions answered by the control group (*C*) showed that 70% had a distinct expectation about the experiment that ran counter to the experimental explanation. Of these subjects, some had a conscious or pre-conscious remembering set, and some believed that their curiosity was being tested. Nine of the ten subjects reported behaviours while viewing the slides that might be considered to reflect 'structure-building', or, at least, behaviours conducive to later remembering.

Three subjects thought the slide procedure was a test of memory (question 2, item 9). These subjects all reported that they tried to memorize visual images of the stimuli (question 3, item 3); and checked one or more additional methods that might aid later recall (items 2 and 4).

Four subjects indicated that they thought the slide procedure was a test of curiosity (question 2, item 6). Although they checked no items that dealt with the expectation of having to recall stimuli, three nevertheless tried to memorize visual images of the slides, and all tried to label them (question 3, items 3 and 4).

Three subjects maintained that their only expectation was what the experi-

menter had told them: namely, that the purpose of the experiment was to time viewing of the stimuli in order to see which were most interesting (question 2, item 1). However, two of these subjects indicated that they labelled the stimuli as they went along; and one of the two, that she tried to memorize the figures visually.

Collectively, these results suggest that efforts to memorize and some kind of 'structure-building' may take place spontaneously, i.e. without the experimenter's encouragement, and across different experimental conditions.

DISCUSSION

The experiments started from the question: Do Berlyne's subjects look longer at the collative stimuli because a curiosity motive has been aroused, or because a remembering set has been evoked? Control group data suggest that both factors are at work in subjects given 'neutral' instructions patterned on Berlyne's. The two experimental conditions, curiosity-arousal and remembering set, work in the direction of separating the two effects.

Although curiosity-arousal and a set to remember increase attention to collative stimuli to about the same overall extent, a difference comes to light on further inspection of Table 1. Ambiguous figures are responded to in approximately equal measure by subjects in all four conditions, but the other two classes of stimuli are responded to differently. Curiosity-arousal particularly prolongs attention to *incongruous* stimuli, whereas an experimentally induced remembering set especially heightens attention to *complex* stimuli. This relation is borne out by data on the control subjects who believed that either their curiosity or their memory was being put to the test. The controls with something akin to a 'curiosity set' viewed the more incongruous stimuli for 61 % of the time, on the average, compared with 53 % for the complex stimuli. By contrast, controls with an implicit remembering set viewed incongruous stimuli 52 % of the time, compared with 59 % for the complex stimuli. (These patterns are similar to those of the curiosity-arousal and remembering groups, respectively, as shown in Table 1.)

One way to test the apparent relations between curiosity-arousal and incongruity and between a remembering orientation and complexity is to ask: how many subjects in these two experimental conditions showed the expected pattern? A 2×2 classification, based on experimental condition (C-A or R) and on which group of stimuli, incongruous or complex, was viewed proportionately longer, reveals a non-significant trend in the expected direction. The main problem in such an analysis, of course, is the small sample number. When, however, n is increased by considering, together with subjects in the C-A condition, control subjects who in fact had a self-induced curiosity set, and by including with subjects in the R group those controls who actually had a remembering set, the degree of relation improves. Fisher's exact test, applied to the resulting data, results in $P < 0{\cdot}05$ (one-tailed).

The possibility that attention to incongruity and complexity may tap different processes is in line with at least two other research findings. In a recent paper Berlyne reports the failure of complexity to trigger epistemic curiosity, but attributes the result to inadequacy of stimulus-materials (Berlyne & Frommer, 1966). He found that children asked more questions about incongruous pictures than non-incongruous ones;

but that amount of information did not affect questioning. (Berlyne had predicted more questions in response to pictures accompanied by a short factual story than to the pictures alone.) Smock & Holt (1962), in a study of children's perceptual curiosity, found an overall preference for collative stimuli: but complexity contributed the smallest amount to the effect, while incongruity attracted by far the greatest amount of differential viewing.

Does complexity *per se* arouse curiosity (Berlyne's position)? Or must the complex stimulus also be incongruous (surprising) in order to trigger curiosity? The latter possibility deserves further investigation. Whichever is the case, it seems that the different types of collative stimuli may not engage exploration for the same reason, even in the simplified laboratory situation used by Berlyne and the present investigators. Caution is needed in using the concept of curiosity motivation to account for attention to collative stimuli, as Cantor *et al.* (1963) has suggested. Studies of perceptual curiosity have not employed criteria independent of the viewing procedure for demonstrating the presence of a curiosity motive: for example, Berlyne (1954*a*, 1957, 1958*a*); or Smock & Holt (1962), who exposed subjects to a curiosity-arousing situation prior to stimulus-viewing, but had no control subjects in a non-arousal condition. The presence of a *curiosity motive* in subjects who are viewing stimuli that vary in collative properties needs to be supported by independent criteria.

Warm thanks are due Norman Livson for a stimulating discussion of the results, and M. Brewster Smith for a critical reading of the manuscript. We also wish to thank students in Psychology 218 who helped to pre-test the stimuli used in these experiments.

References

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Br. J. Psychol.* **41**, 68–80.

Berlyne, D. E. (1954*a*). An experimental study of human curiosity. *Br. J. Psychol.* **45**, 256–65.

Berlyne, D. E. (1954*b*). A theory of human curiosity. *Br. J. Psychol.* **45**, 180–91.

Berlyne, D. E. (1954*c*). Knowledge and stimulus-response psychology. *Psychol. Rev.* **61**, 245–54.

Berlyne, D. E. (1957*a*). Conflict and choice time. *Br. J. Psychol.* **48**, 106–18.

Berlyne, D. E. (1957*b*). Conflict and information-theory variables as determinants of human perceptual curiosity. *J. exp. Psychol.* **53**, 399–404.

Berlyne, D. E. (1958*a*). The influence of complexity and novelty in visual figures on orienting responses. *J. exp. Psychol.* **55**, 289–96.

Berlyne, D. E. (1958*b*). The influence of the albedo and complexity of stimuli on visual fixation in the human infant. *Br. J. Psychol.* **49**, 315–18.

Berlyne, D. E. (1958*c*). Supplementary report: complexity and orienting responses with longer exposures. *J. exp. Psychol.* **56**, 183.

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity.* New York: McGraw-Hill.

Berlyne, D. E. (1963*a*). Motivational problems raised by exploratory and epistemic behavior. In *Psychology: a Study of a Science*, vol. 5, ed. S. Koch. New York: McGraw-Hill.

Berlyne, D. E. (1963*b*). Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. *Can. J. Psychol.* **17**, 274–90.

Berlyne, D. E. & Frommer, F. (1966). Some determinants of the incidence and content of children's questions. *Child Dev.* **37**, 177–89.

Butler, R. A. (1953). Discrimination learning by rhesus monkeys to visual-exploration motivation. *J. comp. physiol. Psychol.* **46**, 95–8.

Cantor, G. N. (1963). Responses of infants and children to complex and novel stimulation. In *Advances in Child Development*, vol. 1, ed. L. Lipsitt & C. Spiker. New York: Academic Press.

CANTOR, G. N., CANTOR, J. & DITRICHS, R. (1963). Observing behavior in pre-school children as a function of stimulus complexity. *Child Dev.* **34**, 683–9.

DEMBER, W. N. (1956). Response by the rat to environmental change. *J. comp. physiol. Psychol.* **49**, 93–5.

DEMBER, W. N. (1961). Alternation behavior. In *Functions of Varied Experience*, ed. D.W. Fiske & S. R. Maddi. Homewood, Ill.: Dorsey.

FANTZ, R. L. (1961). The origins of forms perception. *Scient. Am.* **204**, 66–72.

FISKE, D. W. & MADDI, S. R. (1961). A conceptual framework. In *Functions of Varied Experience*, ed. D. W. Fiske & S. R. Maddi. Homewood, Ill.: Dorsey.

HARLOW, H. F., HARLOW, M. K. & MEYER, D. R. (1950). Learning motivated by a manipulation drive. *J. exp. Psychol.* **40**, 228–34.

HEBB, D. O. & THOMPSON, W. R. (1954). The social significance of animal studies. In *Handbook of Social Psychology*, vol. 1, ed. G. Lindzey, pp. 532–61. Cambridge, Mass.: Addison-Wesley.

LEUBA, C. (1955). Toward some integration of learning theories: the concept of optimal stimulation. *Psychol. Rep.* **1**, 27–33.

MADDI, S. R. Exploratory behavior and variation seeking in man. In *Functions of Varied Experience*, ed. D. W. Fiske & S. R. Maddi. Homewood, Ill.: Dorsey.

MADDI, S. R., PROPST, B. & FELDINGER, I. (1965). Three expressions of the need for variety. *J. Pers.* **33**, 82–98.

MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A. & LOWELL, E. L. (1953). *The Achievement Motive*. New York: Appleton-Century-Crofts.

MENDEL, G. (1965). Children's preferences for differing degrees of novelty. *Child Dev.* **36**, 453–65.

MITTMAN, L. R. & TERRELL, G. (1964). An experimental study of curiosity in children. *Child Dev.* **35**, 851–5.

ORNE, M. T. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Am. Psychol.* **17**, 776–83.

RAPAPORT, D. (1960). On the psychoanalytic theory of motivation. In *Nebraska Symposium on Motivation*, ed. M. K. Jones. Lincoln: University of Nebraska Press.

SMOCK, C. D. & HOLT, B. G. (1962). Children's reactions to novelty: an experimental study of 'curiosity motivation'. *Child Dev.* **33**, 631–42.

WELKER, W. I. (1956*a*). Effects of age and experience on play and exploration among young chimpanzees. *J. comp. physiol. Psychol.* **49**, 223–6.

WELKER, W. I. (1956*b*). Some determinants of play and exploration in chimpanzees. *J. comp. physiol. Psychol.* **49**, 84–9.

(*Manuscript received* 27 *July* 1966; *revised manuscript received* 9 *June* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 387–394
Printed in Great Britain

EFFECT OF DELAY OF ESCAPE REWARD AND RESPONSE COMPATIBILITY ON HUMAN REACTION TIME

By ATHAN KARRAS*
Ohio State University

The gradient of escape reward under compatible and incompatible responses performed during delay of reward by human subjects was studied. A discrimination jump reaction time task was used; escape was from aversive noise. Subjects practised the same jump response and one of two delay responses on pre-reward discrimination trials and were told that the pre-reward jump response was the escape response on noise trials. During noise trials reward followed the escape response by 0, 0·4, 4·0 or 8·0 sec. A no-noise control group for each kind of delay response was included. The jump speeds for both kinds of delay responses were the same for pre-reward trials and also for reward trials for zero delay. For reward trials the delay groups with the incompatible response were slower than the compatible response group counterparts on all but 8·0 sec delay. The no-noise controls did not improve. Extinction was immediate for all reward groups.

The shape of the gradient of primary negative reward for rats performing an instrumental response (e.g. escape from shock) has been found to be similar to the gradient of primary appetitive reward (Fowler & Trapold, 1962). Characteristically for either reward, there is a decrement in instrumental performance which is a negatively accelerated function of the delay interval.

Fowler & Trapold proposed that the decrement for either reward may be accounted for by Spence's (1956) formulation of the role of the inhibiting intervening variable, I_t, in instrumental performance for appetitive rewards. Briefly, competing responses made in the goal area during delay of reward are conditioned to goal cues and become anticipatory by way of the similarity of cues in the alley, thereby interfering with the appropriate response.

However, there is no unequivocal evidence that competing responses during the delay of primary appetitive or aversive reward actually interfere with the appropriate response. With regard to primary appetitive reward, Spence (1956) cited a study by Carlton. In this study the likelihood of incompatible responses was reduced by orientating the subject to a food cup and a response bar to be pressed to obtain food in the cup. This procedure most certainly confounded compatible kinaesthetic cues with visual cues from viewing the cup and manipulandum, located in the same place on all trials. Fowler & Trapold offer only anecdotal evidence for aversive reward. Also, secondary reinforcement from specific place cues was present in their study because the subject traversed the same alley on all trials. Such cues may explain why decrement in speed was gradual, whereas Riesen (1940), who had used a discrimination task in order to minimize the effects of such cues, had found the decrement to be abrupt. Thirdly, there was no evaluation in their study of the effects on drive level of shock adaptation associated with the various delay intervals. Therefore, it is not clear to what the shape and extent of the gradient of delay of escape are due.

The present study investigated the effects of specified compatible and incompatible responses made during various delays in escape reward on human reaction time (RT).

* Now at Hillside Hospital, Glen Oaks, New York.

The method minimized effects of other sources of specific secondary reinforcement by using a discrimination task. Human subjects were used to determine the applicability of the incompatible response hypothesis (*a*) to human subjects receiving primary aversive reward, and (*b*) to human psychomotor performance, which has been suggested by Bilodeau & Bilodeau (1958) and Greenspoon & Foreman (1956).

METHOD

Subjects

The subjects were 154 undergraduates between the ages of seventeen years six months and thirty-eight years. One hundred and twenty of them were used in the main experiment, which was composed of ten groups with twelve subjects per group. The remaining subjects were used in additional control groups after the completion of the main experiment. The description of procedures and results for the additional control groups are last in the method and results sections. All subjects had to tolerate a pre-test white-noise intensity of 99 db, sound pressure.

Apparatus

The reaction time (RT) apparatus had a baseboard and a vertical panel on top of a table. The baseboard had a white start-button for the subject which was 9 in. from the press-buttons of two telegraph keys mounted 9 in. apart. The start-button was also a safety device to prevent a response light from lighting up if the subject released his finger from the start-button before the response light flashed on. In front of the press-buttons of the telegraph keys, 8·5 in., was the vertical panel, in which three 0·5 in. jewelled pilot lights were placed 18 in. above the baseboard and at about eye-level for the subject. One of the lights, the ready signal, was centred and appeared green when illuminated. On either side of the ready signal, 8·5 in., was a red response light. The ready signal was slightly brighter than the response signals. A light pressure on the press-button of the right telegraph key turned off the right response light, and the press-button of the left telegraph the left response light.

The ready light flashed on for 0·3 sec. A response light remained illumined until the subject hit the correct button. Two preparatory intervals (PIs), the interval between the appearance of the ready signal and the onset of a response light, were used (5 and 7 sec). The sequence of the presentation of the PIs and the sequence of right and left response lights were quasi-randomly determined, with the restriction that no interval or response light be presented more than three times in succession. All subjects received the same PI sequence, whereas the sequence for the response lights varied among subjects.

The inter-trial interval commenced, if no sound was presented, when the subject pressed the correct button. If sound was presented, it began with sound offset. In either case, 1 sec later the ready signal flashed.

The white noise, when presented, commenced with the appearance of a response light. Its offset was 0, 0·4, 4·0 or 8·0 sec after the subject responded correctly. The sound came from a Grason–Stadler white-noise generator and was delivered at 99 db, sound pressure. All intervals were timed electrically.

Illumination for the room came from two 15-watt fluorescent tubes in a lamp placed at the back of the apparatus where the experimenter sat.

A standard Electric timer, activated simultaneously with a response light and terminated by the subject's appropriate response, recorded RT to the nearest 0·01 sec.

Procedure

Whenever a response light appeared, the subject made a jump reaction from the start-button to the press-button of the appropriate telegraph key, using the middle finger of his preferred hand. The time to execute this response was the performance measure used.

After being subjected to the sound stimulus which he (she) had to tolerate without experiencing any physical sensations, the subject was then assigned to either an immediate return (IR) (more compatible) or detained return (DR) (less compatible) group. For both groups the experimenter pointed to the start-button on which the subject rested his middle finger and then

to the press-button of the telegraph keys, informing the subject how to put out the red light which followed the green light. Instructions avoided any indication of the desirability of speed.

For the IR group, the experimenter also said: 'Immediately after putting out the light, return to the side of the white button. Rest your hand on the grey platform (the table). The green light will flash on shortly afterwards as a signal to return immediately to the white button for the next trial.' The subject's hand on the table was 1·25 in. lower than the start-button, with the middle finger approximately 8·5 in. from either the start- or the press-button located on the same side as the hand used.

For the DR group, the experimenter said: 'Leave your finger on the telegraph key after putting out the light. The green light will flash on shortly afterwards as a signal to return immediately to the white button for the next trial.'

For both groups final instructions were: 'The sound may come on when the red light flashes on. It will stop sometime before the green light appears, but the sound will not stop if you fail to turn off the red light.'

Practice trials. After these instructions, the experimenter administered fifteen practice trials to the subject without sound and corrected his movements if necessary by reading the appropriate instructions. The subject wore earphones during all trials, regardless of whether sound was presented, and removed them during rest periods which lasted 1 min. between each series of fifteen trials. Before commencing a series of trials, the experimenter told the subject: 'Let's continue as before. Your job is to remember to turn out the red light each time it comes on.'

During the rest period following the first series of trials, the experimenter obtained data on the subject. Four more series of trials without sound followed. Medians of the last and next to last series were calculated. If the subject was slower than 1 sec on the median of the last series, or if he gained more than 0·06 sec or lost more than 0·10 sec on the last median as compared with the next to last median, the experiment was discontinued. Twenty subjects were discarded by these criteria, twelve of them in the IR group. The correction method was used in case the subject struck the wrong press-button. Virtually no incorrect responses (three) were made after the practice trials.

Reinforcement trials. Three series of reinforcement trials followed. Within the IR or DR groups the subject was randomly assigned to one of five treatment conditions, four of which received the sound stimulus, the fifth being a control group performing as before. For the four sound groups, the sound commenced with the response signal and terminated 0, 0·4, 4·0 or 8·0 sec after the subject responded correctly. As in the practice trials, the subject continued to make his detained or immediate return response during the delay of reinforcement.

Extinction trials. Two series of extinction trials followed the rest period after the last series of reinforcement trials, and were identical to the practice trials.

Post-extinction questions. After the extinction trials non-control subjects were asked three questions, answers being recorded verbatim: (1) What do you think you should have been doing on the noise trials? (2) How did the noise affect your speed to react? (3) Who controlled how long the noise lasted? If the answer to (3) was 'the experimenter', the following was also asked: (4) While reacting to the noise, did you remember the instructions that the sound will not stop if you failed to turn off the red light?

Additional control groups

Control for response–response interval. It was possible that increased inter-response interval interacted with high drive to induce performance decrement. Two conditions (n = 5 per group) were used, one of which was another DR 0 group. The second, DR 0–4·0 sec, had immediate reward, but the subject maintained the DR position for an additional 4 sec, as in the DR 4 sec group in the main experiment.

Control for noise adaptation. It was possible that the longer exposure to sound of the longer delay groups decreased the drive value of the noise for these subjects through greater adaptation. Two conditions (n = 12 per group) were used, one of which was another DR 0 sec delay of reward. The second, DR 8–0 sec, had 8·0 sec delay of reward for the first thirty-eight reinforcement trials, followed by fifteen trials with immediate reward without a rest period either after the thirty-eighth trial or any time during these last fifteen trials. The subjects were not informed of the change in delay.

RESULTS

Scores used and preliminary tests

For each subject the median was obtained for each series of fifteen trials, beginning with the last pre-reinforcement series. A speed score was obtained by transforming a median into its reciprocal $\times 10{,}000$. Thus for each subject there were one pre-reinforcement, three reinforcement and two extinction scores.

The score used as the estimate of the subject's final performance for the reinforcement condition was the one obtained from the third (last) series of reinforcement trials. This measure was considered stable, since a non-significant t-value was found in a t-test for paired observations performed between the mean of the reciprocals of the median of the first five trials of the last series vs. the comparable mean of the last five trials of this series for the group with the largest difference between these means (DR, zero delay). (Such a test favours finding a significant difference.)

It was found that the within-subject variances of the IR and DR control groups were on the average half that of the experimental groups. Individual comparisons were therefore first made between 0·0, 0·4, 4·0 and 8·0 delay groups and their appropriate control group, using difference scores obtained by subtracting the pre-reinforcement score from the last reinforcement score. The Dunnett t-statistic was used for these multiple comparisons (Winer, 1962). For IR, all but the 8·0 delay group were significantly different from the IR control group; for DR, the 0·0 and 0·4 delay groups were significantly different from the DR control group. Where necessary, degrees of freedom were halved. The significant differences between some of the experimental groups and their respective control groups justified further statistical tests among the experimental groups.

In the subsequent analyses no differences were found between men and women subjects within groups.

Pre-reinforcement and reinforcement performance

The means and standard deviations of the pre-reinforcement and last reinforcement scores for the groups in the main study are presented in Table 1. (Two subjects were discarded and replaced because of aberrant reinforcement difference scores. One subject in IR 4·0 sec improved 85 units; a subject in DR 8·0 lost 46 units. Both scores are more than 3 S.D. beyond their means, when mean and S.D. for their respective groups include these values. Inclusion of these extreme values would make more striking the present findings between the IR and DR groups.)

A $2 \times 4 \times 2$ analysis of variance with repeated measures on the last factor (Winer, 1962) was made on the results for these groups; the results are given in Table 2. The interaction of interest, $R \times D \times T$, was significant at $P < 0{\cdot}02$.

Individual comparisons within the IR and DR groups were made using the Newman–Keuls method. For the IR group all comparisons, except for 0 vs. 0·4 sec and 4·0 vs. 8·0 sec, were significant at $P < 0{\cdot}01$. For the DR group the 0·0 vs. 0·4 sec and 0·0 vs. 4·0 sec comparisons were significant at $P < 0{\cdot}01$, while the 0·0 vs. 8·0 sec and 0·4 sec comparisons were significant at $P < 0{\cdot}05$. All tests were two-tailed.

For comparisons of the corresponding delays between the IR and DR groups

t-tests were used because of the *a priori* hypothesis that IR would lead to superior performance than DR for the non-zero sec delays. As expected, the zero delay comparison was not significant ($P > 0{\cdot}10$). However, the 0·4 sec and 4·0 sec comparisons were significant ($P < 0{\cdot}05$) but not the 8·0 sec ($P > 0{\cdot}10$). Another comparison was made between IR 0·4 and 4·0 sec combined, and DR 0·4 and 4·0 sec combined; it was significant at $P < 0{\cdot}01$. All tests were two-tailed. The last comparison was made with regard to Grice's (1948) and Riesen's (1940) findings that only the shorter delays (up to 5 sec) show learning when the effects of secondary reinforcers are minimized.

Table 1. *Means and standard deviations of speed scores of pre-reinforcement and reinforcement: trials of groups in the main study*

	Pre-reinforcement		Reinforcement	
	Mean	S.D.	Mean	S.D.
IR group. Delay (sec)				
0·0	147·50	20·34	180·58	24·37
0·4	135·92	22·20	160·83	32·90
4·0	145·08	23·05	154·17	21·53
8·0	137·08	21·05	138·50	21·81
Control	149·25	26·95	149·25	27·90
DR group. Delay (sec)				
0·0	151·08	30·19	190·67	33·36
0·4	154·83	43·82	167·42	41·42
4·0	146·50	21·44	144·00	24·14
8·0	146·50	24·63	155·17	27·31
Control	163·00	32·26	165·42	38·50

Table 2. *Analysis of variance of groups in the main study for pre-reinforcement and reinforcement: trials* (*control groups excluded*)

Source	D.F.	M.S.	*F*
Between subjects			
Response	1	2,394·19	1·63
Delay	3	5,080·98	3·47*
R × D	3	795·87	0·54
Subjects within groups	88	1,465·46	
Within subjects			
Trials	1	12,065·06	116·62**
R × T	1	77·52	0·75
D × T	3	2,810·03	27·16**
R × D × T	3	355·26	3·43*
Trials × subjects within groups	88	103·46	

* $P < 0{\cdot}02$; ** $P < 0{\cdot}001$.

Extinction performance

Extinction scores were analysed separately from reinforcement scores because of their greater variability. With or without the control groups included, F_{max} was always significant at $P < 0{\cdot}05$. Preliminary individual comparisons revealed that no reinforcement group was superior in performance to its control group (all $P > 0{\cdot}10$). No other analyses were therefore made.

Questionnaire response

Question 1. The majority of subjects in each group replied that their task was to turn off the red light. The distribution of subjects who said that they were to respond more quickly because of the sound is as follows: IR 0·0: four subjects; IR 4·0: two subjects; DR 0·0: one subject; DR 0·4: three subjects.

Question 2. Answers to this question were not useful and are omitted.

Question 3. IR 0·0 and 0·4, and DR 0·0 and 0·4 groups had a majority of subjects believe they controlled sound offset. The other delay groups for IR and DR had virtually all subjects claim they believed that either the experimenter or the apparatus controlled sound offset.

Question 4. This question was asked only of those subjects who failed to state that they controlled sound offset and was asked mostly of IR 4·0 and 8·0 and DR 4·0 and 8·0 groups. All but one subject said that they remembered during the reinforcement trials that the noise would not go off if they failed to turn out the red light.

Additional control groups

Response–response interval. The means of the DR 0·0 delay with 1 sec inter-trial interval and DR 0·0 delay with 5 sec inter-trial interval were extremely similar, suggesting that no significant difference would be found with increased sample size.

Noise adaptation. For each group, differences were obtained between the transformed pre-reinforcement score and scores for reinforcement trials 34–38 and 40–44. A 2×2 factorial analysis of variance with repeated measures of these differences showed that prior to the switch-over the 0·0–0·0 delay group was significantly faster ($P < 0{\cdot}005$) than the 8·0–0·0 group. However, within the first five trials after the switch-over there was no significant difference between the two groups ($P > 0{\cdot}10$).

DISCUSSION

The present study confirmed Spence's hypothesis of the inhibiting nature of incompatible responses performed during delay of primary reward on the instrumental response. It also suggests a parallel between appetitive and escape conditioning and that a common principle may underlie human and animal motor performance.

Determination of which delay-of-reward response was more incompatible was made from Smith's (1962) findings that dynamic visual feedback, displaying the subject's movements, leads to better performance than static feedback, displaying the end-result, even if his movements are inverted. Applied to the present study it is seen that the DR response kept the subject in a completed response position, so that his posture and arm were opposed to initiating or executing a jump response. On the other hand, the IR response had the subject in motion or in a state of preparedness, which is much more similar to the performance or the anticipation of the instrumental response. In either case, the cues rewarded became anticipatory through the I_t mechanism, with the DR response interfering more with the instrumental response than the IR response.

The present procedure made two assumptions. First, that the discrimination task

was trivial and that the subject did not rely to any important degree on remaining in the goal area in order to utilize goal cues for learning the discrimination; secondly, that the subject was not confused as to what the appropriate instrumental response was. Support for these assumptions comes from the explicit instructions given to the subject before each series of trials, the lack of performance decrement for the delay groups on reward trials as compared to pre-reward trials and the answers to questions one and four.

It is of interest to note that the incompatible response group in the present study and the subjects in Grice's study had the largest extent of their decrement occur within the first ½ sec of delay of reward, i.e. the shape of the gradient is concave downward. In addition, performance was extremely impaired with 4–5 sec delays in the present study and in Grice's and Riesen's. However, it must be pointed out that Grice's shorter delay groups were actually instances of different lengths of within-chain responding. The rapid decrement, however, is to be expected from Spence's notion that such would be the case when secondary reinforcement is minimized or made irrelevant. The most common source of such reinforcement is from place cues which were made irrelevant in this study by using a discrimination task.

The more gradual decline in speed of the IR group may have been caused by the performance of a preparatory response instead of a movement, coupled with a diminution in the incentive to escape as delay increases. The additional control groups demonstrated that neither increased response–response interval, nor decrease in effective drive played an important role in the present study.

The present findings are not in line with the view of Maltzman & Raskin (1965) that the onset of the orienting reflex (OR) or of heightened arousal has reinforcing effects, independent of OR reduction (offset). However, Berlyne, Borsa, Hamacher & Koenig (1966) did obtain results congruent with this view for paired-associate learning. It is hard at this time to account for these differences in results. Perhaps an attack on the problem could start by noting that in the Berlyne *et al.* study the subject could rehearse the correct response implicitly until the next trial, whereas in the present study the intervention of the delay-of-reward response superseded such activity.

Extinction for all noise groups was rapid. This is consistent with Logan's (1960) and Renner's (1963) findings. Some of the factors which may have contributed to rapid extinction are: (*a*) secondary reinforcement, especially place cues, were at a minimum, (*b*) the inter-response interval during extinction trials was shorter than during reinforcement trials, preventing delay-of-reward subjects from performing in a way similar to the reinforcement trials, and (*c*) the noise was omitted. Thus subjects in any group could readily discriminate the change in stimulus conditions between the reinforcement and extinction trials. Spence (1966) demonstrated that the omission of just the noxious stimulus in conditioning the eyeblink in human subjects is sufficient for extinction to be rapid, provided the nature of the task is not masked.

This investigation was supported in part by institutional grant funds of the Ohio State University Medical Center, Rotary No. 42071, Project No. 285.

REFERENCES

BERLYNE, D. E., BORSA, D. M., HAMACHER, J. H. & KOENIG, I. D. V. (1966). Paired-associate learning and the timing of arousal. *J. exp. Psychol.* **72**, 1–6.

BILODEAU, E. A. & BILODEAU, I. McD. (1958). Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.* **55**, 603–12.

FOWLER, H. & TRAPOLD, M. A. (1962). Escape performance as a function of delay of reinforcement. *J. exp. Psychol.* **63**, 464–7.

GREENSPOON, J. & FOREMAN, S. (1956). Effect of delay of knowledge of results on learning a motor task. *J. exp. Psychol.* **51**, 226–8.

GRICE, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.* **38**, 1–16.

LOGAN, F. A. (1960). *Incentive*. New Haven: Yale University Press.

MALTZMAN, I. & RASKIN, D. C. (1965). Effects of individual differences in the orienting reflex on conditioning and complex processes. *J. exp. res. Pers.* **1**, 1–16.

RENNER, K. E. (1963). Influence of deprivation and availability of goal box cues on the temporal gradient of reinforcement. *J. comp. physiol. Psychol.* **56**, 101–4.

RIESEN, A. H. (1940). Delayed reward in discrimination learning in chimpanzees. *Comp. psychol. Monogr.* **15**, 1–54.

SMITH, K. U. (1962). *Delayed Sensory Feedback and Behavior*. Philadelphia: Saunders.

SPENCE, K. W. (1956). *Behavior Theory and Conditioning*. New Haven: Yale University Press.

SPENCE, K. W. (1966). Cognitive and drive factors in the extinction of the conditioned eyeblink in human subjects. *Psychol. Rev.* **73**, 445–58.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 16 *February* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 395–401
Printed in Great Britain

AN EXAMINATION OF TEST SITUATIONS INVOLVING SUGGESTIBILITY BY INDIRECTION

By G. SINGER
University of Sydney

The hypothesis that the 'progressive lines test' represents a test of 'suggestibility by indirection' was tested in a series of experiments. In Expt I fifty-six subjects who were given no indirection pretraining did not differ from fifty-six subjects who were given this training. In Expt II it was shown that a systematic 'time error' is operative in this test situation and that this effect confounds the measurement of suggestibility by indirection.

This study was concerned with 'suggestibility by indirection' as postulated by Binet (1884*a*, *b*) and developed into what has been regarded as a standard test situation by later investigators (Brown, 1916; Hull & Forster, 1932; Hull, 1933; Eysenck, 1943; Eysenck & Furneaux, 1945; Stukat, 1958). The 'progressive lines test' and the 'progressive weights test' are the basis of this concept of 'suggestibility by indirection'.

In the 'progressive lines test' a series of five lines, which increase in length, is presented to the subject. These are followed by ten lines of equal length presented successively. The task is to compare each line with the previous one and to make a judgement of 'longer', 'shorter' or 'equal'. The assumption is made that the response category 'longer' is the result of a 'suggestion by indirection' contained in the presentation of the first five stimuli (pretraining series). It is argued here that this assumption needs to be tested by means of a control group to whom only the ten objectively equal lines are presented, without the pretraining stimulus series. The use of a control group is essential in order to allow the experimenter to judge how much of the variance is contributed by the suggestion stimuli.

In the case of the progressive lines test without a control group the task is of an unknown level of difficulty. The subject may give responses 'longer' because of an inability to make a correct discrimination. These error responses will be indistinguishable from 'suggestion responses'. The contention that the trait of suggestibility is normally distributed (Brown, 1916; Hull, 1933) would also hold for random error when the difficulty level is such that the probability of a correct response is 0·5. It is also possible that there is an illusion in this test situation for part of the stimulus display. For example, when the response required is a judgement of length for one stimulus which, because of its measured size, is known to the experimenter as the stimulus for a correct response, contextual conditions for another stimulus may give it the same appearance; the result may be large numbers of responses to this stimulus which are then interpreted as responses due to the experimental treatment (suggestion by indirection). The result is a systematic error.

Generally, this lack of control conditions, which leads to the acceptance of 'random error' or 'systematic error' as the experimental response when the difficulty level has not been determined, is due to an *a priori* assumption that a certain stimulus must always lead to a certain predetermined response with only small random error,

e.g. that the presentation of ten objectively equal lines, if correctly judged, must lead to ten 'equal' judgements when no 'set by indirection' has been established. Without actually checking this *a priori* assumption, the progressive lines test is given under 'set' conditions and experimental responses are scored 'as if' the 'equal' assumption was correct.

EXPERIMENT I

This first experiment provided conditions for the experimental control of these factors. In addition, three scoring methods which have been used by other investigators in this situation were compared. The task was based on the progressive lines test as used by previous investigators (Binet, 1884; Brown, 1916; Hull, 1933; Eysenck, 1943; Stukat, 1958; and others). The difference from previous work lay in the methodology employed: it involved control conditions for a perceptual task without indirection, two different types of instructions, and an increasing, as well as a decreasing, series of pretraining stimuli.

Method

Plan of experiment

The factorial design is set out in Table 1. Subjects were allocated to one of the four groups for each stimulus. The main effects were also used to compare three different measures of suggestibility which had been used by previous investigators. Two groups were given 'non-forcing' or implicit instructions and two groups 'forcing' instructions. Within each set of instructions one group was given a pretraining series and the other group no pretraining series. Since each subject was tested under two conditions with a 2 in. and a 6 in. stimulus the result was a split plot $2 \times 2 \times 2$ factorial design.

Table 1. *Factorial design for Expt I*

Stimulus	Pretraining	Instructions	
		Non-forcing	Forcing
6 in.	No pretraining	Control group 1	Control group 2
	Pretraining (ascending series)	Experimental group 1	Experimental group 2
2 in.	No pretraining	Control group 1	Control group 2
	Pretraining (descending series)	Experimental group 1	Experimental group 2

Apparatus

A box $30 \times 24 \times 18$ in. was used to present the lines to the subject. One end of the box was fitted with an opening for viewing where the subject was asked to sit and await instructions. The other end of the box was fitted with a slot for the insertion of the cards. The cards were 18×8 in. matt-finished ivory board; the lines, commencing 6 in. from the bottom of the card, were of $\frac{1}{2}$ in. black Scotch tape. The box was illuminated with two 25 W globes reduced to one-third capacity with a 'Variac' transformer. A decade timer was used to control the exposure time.

Subjects

Subjects were 112 children with an age range from 11 to 15 years, including approximately equal numbers of girls and boys.

Procedure

Subjects were assigned alternately, as they came along, to either control or experimental groups. They were tested individually. The order of presentation of the 2 in. and 6 in. stimulus series was counterbalanced. Instructions were followed immediately by pretraining or by experimental trials, depending on the group (see Table 1). Each card was presented for 1 sec and the changeover of cards took $1\frac{1}{2}$ sec. There was a break of 30 sec between stimulus series.

The two types of instruction and the pretraining procedure were as follows.

Non-forcing instructions. 'It is the purpose of this experiment to test your ability to estimate the length of lines. I am going to show you the first line when I say "ready", and shortly after I will show you a second line. I want you to tell me if the second line that you see is longer or shorter or just the same length as the first one. Then I'll show you the third line and you tell me whether that one is the same length or shorter or longer than the line you saw just before it. . . and so on. Always tell me as soon as I show you the next line whether it is shorter or longer or the same length as the one before it.'

Forcing instructions were the same as the non-forcing instructions, except that the sentence 'All lines are of different length' was inserted after the first sentence of the non-forcing instructions. An attempt was made to emphasize this line consistently.

Responses 'equal' were allowed in all cases.

'*Pretraining*' for the 6 in. stimulus line consisted of the presentation of five lines increasing in length from 2 in. to 6 in. at intervals of 1 in. and training for the 2 in. stimulus line commenced with a 6 in. line followed by four lines decreasing in length at intervals of 1 in. to the 2 in. line. In the experimental trials following each of the training trials the 6 in. or 2 in. lines were presented separately ten times each.

'*No pretraining*' consisted of the presentation of the relevant stimulus line only once, followed by the experimental trials of ten separate presentations of each of the stimulus lines, as for the pretraining groups.

Methods of scoring

Three methods of scoring were used. (1) The number of responses 'longer' made by each subject for the 6 in. stimulus, or of responses 'shorter' for the 2 in. stimulus, over the ten experimental trials. This was the method used by Brown (1916), Hull & Forster (1932) and Hull (1933) and is here called the 'uncorrected set score'. (2) A summation of both 'longer' and 'shorter' responses for the ten experimental trials (for 2 in. and for 6 in. stimuli). This method scored perceptual errors since all the lines were objectively equal. It was used by Eysenck (1943) and Eysenck & Furneaux (1945) and called the 'personal suggestibility score'. (3) The 'shorter' responses for the 6 in. stimulus were subtracted from the 'longer' responses and vice versa for the 2 in. stimulus. This method was also used by Eysenck (1943) and Eysenck & Furneaux (1945) and called the 'impersonal suggestibility score'. For scoring methods 2 and 3, Eysenck's terms 'personal' and 'impersonal' are used here. The different scores are neither mathematically nor experimentally independent. They have been compared because they have been used by previous experimenters. It is possible that one method may be a better indicator of 'suggestibility' than another since it may provide statistical control for some random or systematic perceptual effect.

Results

The averaged results for all four groups under the two conditions of stimulus size for each of the scoring methods are shown in Table 2.

Table 2. *Mean numbers of responses per subject on ten test trials, for all three scoring methods as used in the analyses of variance for both the 6 in. and the 2 in. stimuli*

Scoring method	Stimulus size		C 1	C 2	E 1	E 2
(1)	6 in.	L	2·64	3·43	1·71	3·89
	2 in.	S	1·18	2·75	1·14	3·68
(2)	6 in.	L+S	4·00	6·46	2·39	5·54
	2 in.	L+S	1·65	4·86	1·68	5·36
(3)	6 in.	L−S	1·28	0·39	1·03	2·25
	2 in.	S−L	0·71	0·64	0·61	2·00

C 1, control group 1 (no pretraining, non-forcing instructions); C 2, control group 2 (no pre-training, forcing instructions); E 1, experimental group 1 (pre-training, non-forcing instructions); E 2, experimental group 2 (pretraining, forcing instructions).

Three analyses of variance were carried out. The analysis for the 'uncorrected set score' involving the means of the first two rows of Table 2 is fully shown in Table 3 and serves as an illustration of the analyses for the other two scoring methods.

The interpretation of these results is unequivocal. The analysis of variance in Table 3 shows tht the effects of instructions and of stimulus size were both significant ($P < 0{\cdot}001$). Pretraining, the variable on which the earlier investigators' suggestibility hypothesis was based, did not reach a satisfactory significance level; in fact, when the means for the implicit instruction group for both stimuli were compared, the control groups with no pretraining showed a higher mean value for the response measure which was to be maximized through pretraining. This trend was reversed for the

Table 3. *Analysis of variance for scoring method* (1), '*uncorrected set score*' (*see rows* 1 *and* 2 *of Table* 2 *for group means*)

Source	S.S.	D.F.	M.S.	F	P
Between subjects					
A Instructions	175·0	1	175·0	22·7	< 0·001
C Pretraining	0·6	1	0·6	0·78	N.S.
A × C Interaction	19·5	1	19·5	2·95	N.S.
Error (between subjects)†	830·3	108	7·7		
Within subjects					
B Stimulus size	30·0	1	30·0	11·5	< 0·001
A × B Interaction	4·6	1	4·6	1·77	N.S.
B × C Interaction	6·5	1	6·5	2·5	N.S.
A × B × C Interaction	0·6	1	0·6	0·23	
Error (within subjects)‡	280·3	108	2·6		
Total	1347·4	223			

† Error term for A, C and A × C.
‡ Error term for B, A × B, B × C and A × B × C.

non-directive instruction. It appears that the hypothesis of 'random error' cannot be rejected, and that 'systematic error' is possible, judging by the significant effect of 'stimulus size'. The fact that explicit instructions produced a significantly greater effect than the 'implicit' instructions indicates that the suggestion in the latter case may have been ineffective for some subjects. The use of this test situation (by Brown, 1916, and others, cited above) and the use of this particular scoring method seem to confound a combination of random and systematic errors and 'no effect' of the suggestion. The 'personal suggestibility score' analysis of variance involving the means of rows 3 and 4 of Table 2 shows the same pattern as the uncorrected set score analysis and therefore leads to the same interpretation.

The 'impersonal suggestibility score' analysis of variance involving the means of rows 5 and 6 of Table 2 gave $F = 14{\cdot}9$ (D.F. 1, 108; $P < 0{\cdot}001$) for the pretraining variable, and $F = 5{\cdot}66$ (D.F. 1, 108; $P < 0{\cdot}05$) for the interaction between instructions and pretraining variables; for stimulus size $F = 6{\cdot}24$ (D.F. 1, 108; $P < 0{\cdot}05$). It is the only scoring method which shows a significant effect for the pretraining variable. It appears that this scoring method, which was used by Eysenck, will maximize the effect of suggestion by indirection. The fact that error within subjects

was almost equal in size to error between subjects indicated, however, that this was not a consistent measure for an individual. The 'impersonal suggestibility score' showed no significant effect for instructions. This result was to be expected since the use of this scoring method involved a correction for the total number of error responses.

Inspection of the data for the control groups suggested a predominance of responses 'longer' over 'shorter' (see Table 4). Analysis of variance (McNemar, 1957, case 17) of the data for the 6 in. stimulus gave $F = 5{\cdot}66$ (D.F. 1, 54; $P < 0{\cdot}05$), and of the data for the 2 in. stimulus gave $F = 3{\cdot}33$ (D.F. 1, 54; $0{\cdot}1 > P > 0{\cdot}05$). Hence a predominance of responses 'longer', without any suggestion by indirection, was present in the results for the 6 in. stimulus, but just failed to reach significance in those for the 2 in. stimulus. In all cases, however, the differences between means were in the same direction.

Table 4. *Mean numbers of responses, 'longer' (L) and 'shorter' (S) separately, per subject over ten trials for the two control groups*

Stimulus	Group	L	S	Difference
6 in.	C 1	2·64	1·36	+1·28
	C 2	3·43	3·03	+0·43
2 in.	C 1	1·89	1·18	+0·71
	C 2	3·39	2·75	+0·64

The results for the control groups therefore imply a source of systematic perceptual error. Errors of this kind have been described elsewhere as 'time errors' (e.g. Woodrow, 1951). Interpreted in this way, both control groups tended to show 'positive time errors' for both stimuli.

Experiment II

The findings of Expt I were cross-validated on a second group of subjects since previous observations of time error have involved different experimental conditions. These subjects were given the control conditions of the 'progressive lines test'.

Method

Apparatus. The apparatus used was that of Expt I.

Subjects. The thirty men and women subjects were drawn from an introductory course in psychology.

Procedure. The procedure was that described for control subjects given 'forcing' instructions in Expt I (group C 2). All subjects were tested with the 6 in. and 2 in. stimulus figure; order of presentation was counterbalanced between subjects.

Results

The results and their significance levels are shown in Table 5. They confirm in each instance the operation of a systematic effect. Subjects reported significantly more often that the second line on a successive comparison task was longer than the one seen immediately before, when the time interval between presentation of stimuli was approximately 1 sec and the two stimuli were objectively equal (in fact the same stimulus card was used on all ten trials).

This 'time error' must mask responses, especially when the same response is measured as a manifestation of suggestibility as in the procedure of the progressive lines test (Binet, 1884*a*, *b*; Brown, 1916; Hull, 1933). It is of interest that psychologists interested in perception reported this phenomenon for successive comparisons in all sense modalities, including a weight-lifting test similar to the progressive weights test. Prediction of this systematic error for a particular test situation was complicated by the fact that the time error may be systematically positive or negative in different experimental conditions.

Table 5. *Mean numbers of 'longer' (L) and 'shorter' (S) responses per ten trials per subject in Expt II (experimental conditions as for group C 2 in Expt I) and their comparison for each size of stimulus*

		Mean no. responses			
n	Stimulus	L	S	*t*	*P*
30	6 in.	3·53	2·20	2·89	< 0·01
30	2 in.	3·47	2·07	2·59	< 0·02

DISCUSSION

The two experiments have shown that measurements of individual differences by the 'objective' test method of indirection are influenced by such situational variables as form of instructions and stimulus size. Two of three methods of scoring subjects' responses, however, showed no reliable difference between 'pretraining (indirection)' and 'no pretraining (control)': there was no significant increase in the numbers of responses, either 'longer' or 'shorter', which were expected to be increased by pretraining (uncorrected set score), nor any significant difference for the sum of 'longer' and 'shorter' responses for each subject ('personal suggestibility score', Eysenck, 1943); the exception was the third scoring method, which involved taking the difference in number of 'longer' and 'shorter' responses for each subject and so involved a correction for perceptual error ('impersonal suggestibility score', Eysenck, 1943). The case for this score as a measure of suggestibility is weakened, however, by the finding of low, non-significant, internal consistency for the method. The analysis of the data suggests that measurements of suggestibility by indirection have been confounded with 'time errors', as has been confirmed by further analyses of the results for the two control groups of the first experiment and by a second cross-validation experiment under control conditions. In testing situations of this type, successive comparisons must either be avoided or allowance for time errors must be made through the use of control groups.

REFERENCES

BINET, A. (1884*a*). L'hallucination. I. Recherches théoretiques. *Rev. phil.* **17**, 377–412.

BINET, A. (1884*b*). L'hallucination. II. Recherches expérimentales. *Rev. phil.* **17**, 473–502.

BROWN, W. (1916). Individual and sex differences in suggestibility. *Univ. Calif. Publ. Psychol.* **2**, 291–430.

EYSENCK, H. J. (1943). Susceptibility to a visual illusion as related to primary and secondary suggestibility and other functions. *Br. J. Psychol.* **34**, 32–4.

EYSENCK, H. J. & FURNEAUX, W. W. (1945). Primary and secondary suggestibility: an experimental and statistical study. *J. exp. Psychol.* **35**, 485–503.

HULL, C. L. (1933). *Hypnosis and Suggestibility*. New York: Appleton-Century-Crofts.

HULL, C. L. & FORSTER, M. C. (1932). Hallucination and perseverational characteristics of two forms of indirect suggestion. *J. exp. Psychol.* **15**, 700–15.

MCNEMAR, Q. (1957). *Psychological Statistics*. New York: Wiley.

STUKAT, C. C. (1958). *Suggestibility: a Factorial and Experimental Study*. Stockholm: Almqvist & Wiksell.

WOODROW, H. (1951). In *Handbook of Experimental Psychology*, ed. S. S. Stevens, pp. 1224–36. New York: Wiley.

(*Manuscript received* 27 *February* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, 403–412
Printed in Great Britain

A COMPARISON OF VISUAL AND PSEUDO-RECOGNITION THRESHOLDS AS A FUNCTION OF FREQUENCY AND RECENCY

By SANDRA PYKE, NEIL McK. AGNEW and VERA PEZER
York University, Toronto 12, *Ontario*

A series of experiments was conducted to investigate the effects of recency, duration of the interval following recency treatment, and frequency on recognition and pseudo-recognition thresholds. Pre-exposed stimuli were found to have significantly lower thresholds than those for material not pre-exposed. There was no differential effect on performance with intervals of 30 and 60 sec between pre-exposure and threshold determination. However, significantly lower thresholds were obtained with an interval of 30 sec as compared with an interval of 180 sec. Low-frequency words were more influenced by recency of experience than high-frequency stimuli. The effect of frequency on response bias was negligible in the pseudo-perceptual situation and the trend favoured stronger response bias with low-frequency material. The implications of the results are discussed.

The experiments discussed in this paper are designed to test the main and interaction effects of two variables—frequency and recency—on perceptual recognition and on response biasing. That frequency is a major source of variance in visual threshold determinations has been well established. Lower thresholds are typically obtained with high-frequency stimuli as compared with less common stimuli (Engler & Freeman, 1956; Haber, 1965; Hershenson & Haber, 1965; Howes & Solomon, 1951; Postman & Schneider, 1951; Solomon & Howes, 1951; Winnick & Kressel, 1965). When frequency is experimentally manipulated (Solomon & Postman, 1952) similar results have been obtained. Indirect indices of frequency such as familiarity (Daston, 1957) and meaningfulness (Hershenson & Haber, 1965; Kristofferson, 1957) have generally produced the same results as the frequency variable.

The issue of whether frequency serves only to bias responses or does affect the perceptual clarity with which a stimulus is seen has been empirically examined in several recent studies in which the pseudo-perceptual technique is employed. This type of approach is similar to the traditional threshold determination procedures with the major distinction being that no stimulus is displayed. Goldiamond & Hawkins (1958) demonstrated that the number of times a response is emitted in this situation is a function of the frequency of prior experience with the stimulus. Zajonc & Nieuwenhuyse (1964) have also shown lower thresholds for high-frequency material with the pseudo-perceptual method. These pseudo-perceptual studies would seem to support the contention that the frequency effect is not necessarily related to perceptual processes but rather may be due, in part at least, to the subjects' tendency to respond by guessing high-frequency words.

Nevertheless, there is evidence that the frequency effect is not simple. In a pseudo-recognition study it was found (Agnew, Pyke & Pezer, 1968) that the effects of frequency are conditional on the delay between pre-exposure and pseudo-threshold determination, and further that under some delays the low-frequency words obtained significantly lower thresholds. In this instance, the effects were interpreted in terms of the novelty value of the low-frequency material.

While the effects of recency of exposure are well established in studies of memory, recency effects have been relatively ignored in word-recognition studies. There are various methods of experimentally manipulating recency using the traditional word-recognition test procedure. One tactic employed in three studies involved the presentation of anagram tasks prior to threshold determination of the anagram solution words (Eriksen & Browne, 1956; Miller, 1954; Postman & Solomon, 1950). In these studies the anagram solution words had reduced thresholds as compared with control stimuli, but unfortunately the success or failure variable on the anagram words was confounded with recency. A more direct method of manipulating recency is merely to display the stimulus at some point prior to threshold determination and there is some evidence that such a procedure leads to lower thresholds for the pre-exposed stimuli as compared with stimuli not pre-exposed in the experimental session (Neisser, 1954; Smith, 1965*a*, *b*). Thus, while there is some positive evidence for the effects of recency, the conditions under which it operates are not nearly so well documented as in the case of frequency—for example, no parametric studies of recency could be found. Also, the possible interactions between frequency and recency remain to be explored both under perceptual and pseudo-perceptual conditions.

Owing to the relative lack of data it is difficult to predict how frequency and recency variables might interact in either the recognition or the pseudo-recognition situation. In the case of pseudo-recognition, since we have some data (Agnew *et al.*, 1968) suggesting that under recency conditions the effects of frequency are not only reduced but are reversed with low-frequency material showing lower thresholds, we will make the same prediction in the present study. In the traditional word-recognition situation, there is some evidence to suggest that pre-exposure of words will lead to reduced threshold differences between high- and low-frequency words. Haber (1965) and, indirectly, Pierce (1963) present data suggesting this result. However, the interval between pre-exposure and test was not systematically manipulated within the experimental session. With so little data, predictions must remain vague concerning the form of the relationship between frequency and recency.

The purpose of the present experiments was to investigate the main and interaction effects of frequency and several levels of recency in a stimulus-present and in a stimulus-absent situation (pseudo-recognition). As noted in the previous discussion, recency can have two meanings: (1) a situational recency in which some words are pre-exposed in the experimental session and others are not; and (2) a delay-interval recency in which all words are pre-exposed in the experimental situation but at varying time intervals prior to the threshold determination.

METHOD

Subjects

The subjects for all experiments were students between the ages of eighteen and thirty years, enrolled in an introductory psychology course at the University of Saskatchewan. The sample for each experiment was divided into two groups. In all studies, one group was tested on high- and medium-frequency words and the other on low-frequency words. Thirty subjects, seven males and eight females in each group, were involved in each experiment (Expts I, II and III).

Apparatus

A Pharmaceutical Research three-channel tachistoscope was employed for all experiments. One channel served as the pre-exposure field. Two horizontal lines $\frac{7}{16}$ in. apart were placed in this channel as a focus point and all stimuli were presented between these two lines. The sequence of presentation was—first channel: focal lines; second channel: context stimuli for 15 sec; first channel: focal lines for the appropriate delay interval; and second channel: test stimuli. The illumination levels for the pre-exposure (first channel) and exposure (second channel) fields were 32 and 40 ft.-candles respectively.

Three Hunter timers were employed: one to control duration of exposure of the context stimuli (15 sec), one to govern the delay interval between exposures of context and test stimuli, and the third to manipulate exposure durations for threshold determinations of test stimuli. Thus, when the experimenter set the appropriate delay intervals on one timer and pressed a foot-switch, the context stimuli were exposed for 15 sec, the test stimulus dropped into position but was not shown, there was a delay of 30 sec, 60 sec or 180 sec, then the test stimulus was exposed for 0·02 sec. Thresholds were then obtained by manipulating the dial on the appropriate Hunter timer and pressing a hand-switch.

Apparatus used for a task interpolated between pre-exposure of the context stimuli and threshold determination of the test stimulus included a Tandberg tape-recorder, Philips earphones and a tape-recording of the Maudsley Vigilance Test.

Materials

Thirty ten-word sentences (twenty experimental and ten control), each containing one high-frequency and one low-frequency critical word, formed the stimuli for these experiments. Four practice sentences, each containing a high- and low-frequency word, were similarly prepared.

Table 1. *List of critical words*

High frequency		Low frequency	
Journal (100+)	Orchard (20)	Amnesia (5)	Incubus (11)
History (100+)	Lantern (19)	Blarney (14)	Halyard (15)
Country (100+)	Machine (100+)	Debacle (10)	Syringe (15)
Militia (15)	Slavery (21)	Regalia (16)	Padrone (5)
Expense (50–99)	Whistle (50–99)	Salsify (7)	Grapnel (6)
Article (100+)	Ability (49)	Verbose (9)	Euphony (5)
Cabinet (26)	Revenge (29)	Baldric (12)	Inanity (6)
Example (100+)	Pilgrim (14)	Atavism (5)	Hospice (11)
Garland (11)	Pursuit (21)	Czarina (12)	Forager (15)
Quality (50–99)	Kitchen (100+)	Evangel (5)	Liqueur (12)

The critical stimuli were all seven-letter nouns without any double letters. The position of critical words within the sentences was not varied systematically but there were no marked differences between positions for high- and low-frequency words. The critical words themselves were not selected by means of a strict randomization procedure. Instead some attempt was made to have each critical word begin with a different letter of the alphabet, keeping other restrictions in mind. The frequency count for low-frequency words was at least five and not more than sixteen occurrences in 4,000,000 words of test, and for the high-frequency words, from 11 up to over 100 occurrences per 1,000,000 words (Thorndike & Lorge, 1944). These words are presented in Table 1 and the frequency count for each word is in parentheses immediately after the word.

In Expt III the test stimulus or 'pseudo-stimulus' was not a critical word as above, but rather a portion of a Rorschach ink blot, and this was exposed on every trial for 0·02 sec.

Each context sentence was typed in capital letters on a 5 × 7 in. white index card. The forty critical words were typed in the same way on identical cards, as were the critical words for practice trials.

Procedure

For all experiments the subjects were tested individually in sessions ranging from 50 to 90 min and testing of subjects in the high- and low-frequency groups was counterbalanced. All subjects were tested under two delay conditions, 30 sec and 60 sec (Expt I) or 30 sec and 180 sec (Expts II and III). These delays refer to the time interval between presentation of a context sentence and threshold (Expts I and II) or pseudo-threshold (Expt III) determination of a critical word. Each subject also served under recency and non-recency conditions. In the case of the former, an experimental context sentence was exposed, followed by threshold or pseudo-threshold determination of a critical word found within that context. In the case of the non-recency condition, a control sentence was displayed, followed by threshold or pseudo-threshold determination for a critical word from among the experimental sentences. Words from control sentences were never used for threshold or pseudo-threshold determination.

For each subject the ten sentences forming the recency condition (from among the twenty experimental contexts) were determined randomly. Prior to testing each subject, the experimenter shuffled the pack of cards containing the experimental contexts and also the pack of ten cards containing the control sentences. After shuffling, the experimenter placed the pack of critical words in the appropriate order to match that of the experimental contexts. With this procedure, any critical word was presented at any point during the session, thus minimizing consistent position effects for particular words. The pack-shuffling randomization procedure also controlled for potential differences between critical words in that the same word served in both the recency and non-recency conditions for different subjects. Thus any ten critical words from the experimental sentences were used in the recency condition and any ten were used in the non-recency condition.

The order of delay (S or L) and recency (R or NR) conditions was also randomly determined, with the only restriction being that each combination (RS, RL, NRS, NRL) appears five times. Two subjects in each experiment (one in the high-frequency group and one in the low-frequency group) were tested with each order. With this type of design, practice effects were equally spread over all conditions.

The contexts were displayed for 15 sec and the subject was required to read the sentence aloud. If the subject missed some of the context, the experimenter read aloud the relevant portion. During the delay interval the Maudsley Vigilance Test was played and the subject was instructed to keep track of the number of threes that he heard. Following the appropriate delay interval, the threshold or pseudo-threshold for the critical word was determined. In Expts I and II, exposure durations began at 0·02 sec and increased in steps of 0·02 sec up to 0·10 sec and from 0·10 sec at intervals of 0·05 sec until the subject correctly identified the word or until 0·60 sec was reached (trial 15). For Expt III a constant exposure duration of 0·02 sec was employed and the test stimulus was part of a Rorschach blot. Following the procedure introduced by Goldiamond & Hawkins (1958), the subject's responses were recorded on a specially prepared data sheet. The data sheet had twenty columns and fifteen rows and each column was headed by a different word (either high or low frequency, depending on the group). When the subject said the word at the top of the column in which the experimenter was recording, he was told he was correct and the next word in the series was attempted. A pseudo-threshold for a word was terminated either when the subject 'correctly identified' the word or when fifteen flashes of the blot had been presented. The order in which words headed the columns was randomly determined for every subject as was done in Expts I and II.

RESULTS

The number of trials to criterion were totalled for each subject for the five words tested under each recency and delay condition (RS, RL, NRS, NRL). This scoring procedure was followed for all experiments. These four scores for each study were used in separate three-way analyses of variance (Winer, 1962), designed for cases where two variables are of the dependent type. The results of these analyses are presented in Table 2.

Word frequency was found to have a significant effect on thresholds in Expts I and II with high-frequency words having lower thresholds than low-frequency words. In Expt III the frequency effect was not significant, but the trend is in the opposite

Table 2. *Analyses of variance for Expts I, II and III*

		Expt I		Expt II		Expt III	
Source	D.F.	M.S.	*F*	M.S.	*F*	M.S.	*F*
Between-subjects	29						
A (frequency)	1	17376·1	29·0**	6690·1	10·1**	116·0	< 1
Error	28	600·0	—	661·0	—	348·3	—
Within-subjects	90						
B (recency)	1	17184·1	106·8**	14520·0	174·1***	65333·3	187·6***
A × B	1	5044·0	31·3**	1825·2	21·9***	116·0	< 1
Error	28	160·9	—	83·4	—	348·3	—
C (delay)	1	0·3	< 1	691·2	13·1**	396·0	4·0
A × C	1	13·3	< 1	322·7	6·1*	43·2	< 1
Error	28	25·2	—	52·7	—	99·5	—
B × C	1	3·3	< 1	650·8	9·9**	396·0	4·0
A × B × C	1	5·6	< 1	134·8	2·1	43·2	< 1
Error	28	26·6	—	65·9	—	99·5	—
Total	119						

* $P < 0·05$; ** $P < 0·01$; *** $P < 0·001$.

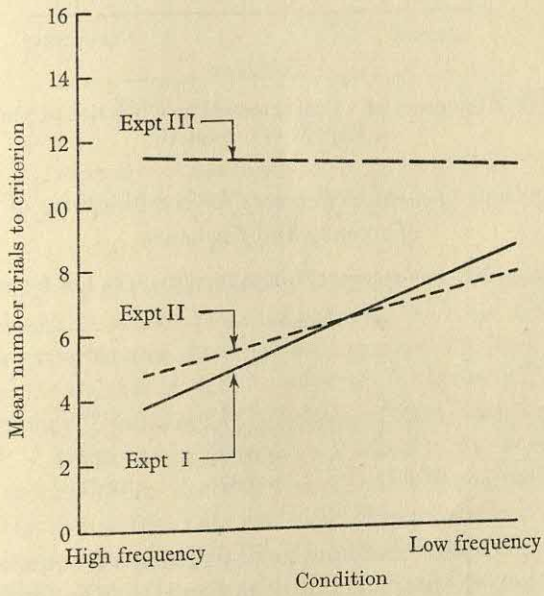


Fig. 1. Trials to criterion in three experiments as a function of frequency.

direction. That is, low-frequency words have lower pseudo-thresholds than do high-frequency words. The frequency effect for all three experiments is depicted in Fig. 1.

In all experiments the effect of pre-exposure of material for subsequent testing (recency) led to significantly lower thresholds as compared with thresholds for stimuli not so recently experienced.

A delay difference of 150 sec was sufficient to affect significantly thresholds in Expt II and this effect approached significance ($P < 0{\cdot}10$) in Expt III. In both cases lower thresholds were found with the short delay condition (30 sec) as compared with the long delay condition (180 sec). However, in Expt I there was no difference in thresholds as a function of delays of 30 sec versus 60 sec.

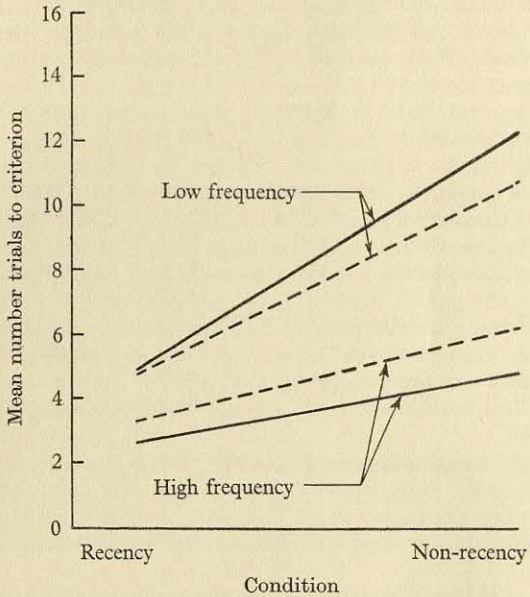


Fig. 2. Effect of frequency on visual threshold as a function of pre-exposure. —, Expt I; - - -, Expt II.

Table 3. *Comparisons of mean differences for combinations of conditions of recency and frequency*

(R, recency; NR, non-recency; H, high frequency; L, low frequency.)

	Expt I		Expt II	
Condition	t	P	t	P
R (H *v.* L)	2·10	< 0·05	1·53	N.S.
NR (H *v.* L)	7·77	< 0·001	4·32	< 0·001
H (R *v.* NR)	4·84	< 0·001	5·99	< 0·001
L (R *v.* NR)	9·12	< 0·001	12·71	< 0 001

The only other significant effect found in Expt I was a frequency by recency interaction. This interaction was also significant in Expt II. The form of this interaction for both experiments is illustrated in Fig. 2. Although both high- and low-frequency words had lower thresholds under the recency condition as compared with the non-recency condition, low-frequency words benefited more by recency treatment than did high-frequency words. In order to determine the source of the interactions in the two experiments, the t-tests of Table 3 were calculated; their magnitudes support the contention that the major effect of recency is with low-frequency material.

Further t-tests were computed to explore the nature of the significant frequency

by delay interaction obtained in Expt II. There was no difference between short and long delay conditions with high-frequency material ($t = 1{\cdot}46$; D.F. 14), while for low-frequency stimuli the difference between delay conditions was reliable ($t = 3{\cdot}21$; D.F. 14; $P < 0{\cdot}01$). Thus the major effect of delay is with low-frequency material. The difference between the two frequency levels was significant under both delay intervals (short delay: $t = 2{\cdot}81$; D.F. 28; $P < 0{\cdot}01$; long delay: $t = 3{\cdot}30$; D.F. 28; $P < 0{\cdot}01$).

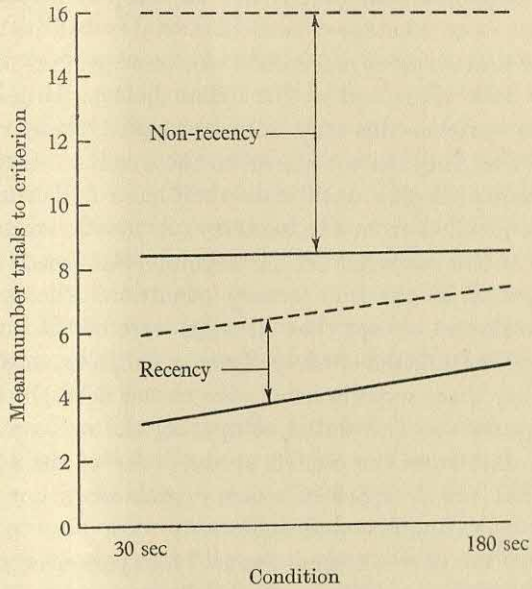


Fig. 3. Interaction of delay and recency in Expts II and III.
—, Expt II; - - -, Expt III.

Although the recency by delay interaction does not reach significance ($P < 0{\cdot}1$) in Expt III, the trend of the results is similar to that of Expt II. Fig. 3 depicts this interaction for both experiments. As expected, in Expt II delay had an effect only with the pre-exposed material ($t = 4{\cdot}13$; D.F. 29; $P < 0{\cdot}001$). The difference between recency and non-recency conditions is significant with both delays (short delay: $t = 9{\cdot}10$; D.F. 29; $P < 0{\cdot}001$; long delay: $t = 7{\cdot}21$; D.F. 29; $P < 0{\cdot}001$).

With Expt III, scores under the non-recency condition, regardless of delay or frequency, were all the same (16). This was to be expected since, as no stimulus word was tachistoscopically presented and the relevant word was not pre-exposed, subjects were guessing at random. Thus the assumption of homogeneity of variance is violated in the analysis of variance reported in Table 2. The data for this experiment were therefore re-analysed with the non-recency condition omitted. With the revised analysis, a significant main effect for delay was observed. No other effects reached significance.

In response to a brief post-experimental questionnaire, none of the subjects in Expt III indicated any awareness that no stimulus word was being presented in the tachistoscope. A majority of subjects in all three experiments stated that the context

sentence was used because it sometimes contained a word which would have to be identified later. Thirty per cent of the subjects in Expts II and III also recognized that this relationship held approximately 50 % of the time.

DISCUSSION

As predicted, recently exposed words in all three studies were found to have significantly lower thresholds than words not as recently experienced. Since the subjects were unaware as to whether a particular trial would be of the non-recency or recency type, their best strategy would be to respond with words from the target sentence which had been presented a short time before. If subjects respond by guessing words in the sentence this may militate against them in the non-recency condition. It may be that they do not attend to the available fragmentary information in the tachistoscopic display until after they have exhausted all the possible response alternatives recalled from the sentence. A fourth control study was conducted to determine if the subjects' set to respond with words from the sentence served to raise thresholds in the non-recency condition. The experiment was procedurally identical to Expt I except that all trials were of the non-recency variety. The recency condition was omitted entirely. Twenty subjects, selected from the same population as the other three experiments, were tested with the same forty critical words. A series of *t*-tests was calculated comparing the non-recency conditions of Expts I and II with data from the control study. None of the *t*-values reach significance, indicating that the presence of recency trials does not artifactually raise thresholds for non-recent stimulus conditions.

The question of how recently a stimulus must be experienced before it will affect recognition or identification speed has been partially answered. There is no difference in performance when material is experienced 60 sec before testing as compared with 30 sec. However, even a short time interval of 180 sec between pre-exposure and threshold determination was sufficient to have an effect on speed of recognition, particularly for low-frequency material. A series of parametric studies is required in order to map out optimal delay intervals under a variety of experimental conditions.

It is demonstrated by the results of Expts I and II that recency as defined is almost as strong a source of variance factor in perceptual recognition as is frequency as defined. The frequency by recency interactions in Fig. 2 illustrate this point. For Expt II, in fact, there is no difference between high- and low-frequency words when both have been recently experienced. This result supports the findings of Pierce (1963) cited earlier, indicating that under recency (reduced population size) conditions the advantage of high-frequency material disappears. It will be recalled also that low-frequency recently experienced words have lower thresholds than high-frequency words not so recently experienced. For example, in Expt I recent low-frequency words have a mean of 4·93 and non-recent high-frequency words a mean of 4·91. Similarly, in Expt II the recent low-frequency mean is 4·81 as compared with the mean threshold of 6·22 for high-frequency words not recently experienced. These findings suggest that a relatively brief exposure to a low-frequency stimulus increases the response strength of this stimulus up to that of a high-frequency stimulus not so recently experienced. This effect may be due in part to the novelty value of

low-frequency material, as has been proposed by the authors in an earlier paper (Agnew *et al.*, 1968). Under the conditions of these experiments, recent experience with a stimulus is just as beneficial in increasing recognition speed as is frequent but less recent experience.

Normally, a stimulus which is experienced frequently is also one which has been experienced recently. Often in perception research, these factors are confounded and it is possible that what has been attributed to frequency may in fact be due, at least in part, to recency.

A comparison of Expts II and III is of major interest because of the markedly different effects of the frequency variable in these two very similar studies. In Expt III not only do high-frequency words not have significantly lower thresholds than low-frequency words, but the trend is in the opposite direction! This finding is in accord with that obtained by Agnew *et al.* (1968) in a study which involved the pseudo-perceptual technique but was otherwise quite unlike that of Expt III. It is argued that the novelty of low-frequency material aids recall perhaps by focusing attention on this unfamiliar material and thereby strengthening responses involving these words to the extent that they are emitted before high-frequency material. In the case of Expts I and II stimulus information from the tachistoscopic display plays a role in further channelling the response and this fragmentary information may counteract the strength of the novelty factor.

The trend for lower pseudo-thresholds with low-frequency material is contrary to the findings of Goldiamond & Hawkins (1958) and Zajonc & Nieuwenhuyse (1964). This contradiction may be due to procedural and design differences. In both the studies cited the stimuli were unfamiliar and meaningless (nonsense syllables and Turkish words respectively) and frequency was defined according to the number of presentations of the various stimuli. Furthermore, in neither of these studies was the degree of recency manipulated.

In other respects the results from Expts II and III are similar but those of Expt III appear to be at a reduced level of intensity. In both studies, lower thresholds are obtained under short delay. Low-frequency words were more influenced by delay and they benefited slightly more by recency treatment than did high-frequency words.

A consideration of these studies points up clearly the complexity of perceptual recognition. A response in these situations is not solely a function of frequency of this response but is also dependent on how recently this response was made and likely also dependent on the novelty of the response.

Questions concerning the relationship between recency and frequency in perception overlap with questions concerning short-term and long-term memory, and, more broadly, with questions of information storage and retrieval. Perhaps the time is at hand for a mutually beneficial methodological and conceptual interchange between the 'perception' and 'memory' fields. For example, eventually, decay and/or interferences curves should be plotted similar to those for traditional memory studies, but using perceptual threshold and pseudo-threshold procedures.

This investigation was supported by grant no. APT-60 from the National Research Council of Canada.

REFERENCES

AGNEW, N. McK., PYKE, S. & PEZER, V. (1968). Frequency and recency effects on response biasing in a pseudo-perceptual task. *Q. J. exp. Psychol.* (in press).

DASTON, P. G. (1957). Perception of idiosyncratically familiar words. *Percept. mot. Skills* **7**, 3–6.

ENGLER, J. & FREEMAN, J. T. (1956). Perceptual behaviour as related to factors of associative and drive strength. *J. exp. Psychol.* **51**, 399–404.

ERIKSEN, C. W. & BROWNE, C. T. (1956). An experimental and theoretical analysis of perceptual defence. *J. abnorm. soc. Psychol.* **52**, 224–30.

GOLDIAMOND, I. & HAWKINS, W. F. (1958). Vexierversuch: The log relationship between word-frequency and recognition obtained in the absence of stimulus words. *J. exp. Psychol.* **56**, 457–63.

HABER, R. N. (1965). Effect of prior knowledge of the stimulus on word recognition processes. *J. exp. Psychol.* **69**, 282–6.

HERSHENSON, M. & HABER, R. N. (1965). The role of meaning in the perception of briefly exposed words. *Can. J. Psychol.* **19**, 42–6.

HOWES, D. H. & SOLOMON, R. L. (1951). Visual duration threshold as a function of word probability. *J. exp. Psychol.* **41**, 401–10.

KRISTOFFERSON, A. B. (1957). Word recognition, meaningfulness and familiarity. *Percept. mot. Skills* **7**, 219–20.

MILLER, C. (1954). Consistency of cognitive behaviour as a function of personality characteristics. *J. Pers.* **23**, 233–49.

NEISSER, U. (1954). An experimental distinction between perceptual process and verbal response. *J. exp. Psychol.* **47**, 399–402.

PEZER, V. & AGNEW, N. McK. (1965). Drive by set interactions in a word recognition task. *Psychonom. Sci.* **2**, 375–6.

PIERCE, J. (1963). Some sources of artifact in studies of the tachistoscopic perception of words. *J. exp. Psychol.* **66**, 363–70.

POSTMAN, L. & SCHNEIDER, B. H. (1951). Personal values, visual recognition and recall. *Psychol. Rev.* **58**, 271–84.

POSTMAN, L. & SOLOMON, R. L. (1950). Perceptual sensitivity to completed and uncompleted tasks. *J. Pers.* **18**, 347–57.

SMITH, W. M. (1965*a*). Visual recognition: Facilitation of seeing by saying. *Psychonom. Sci.* **2**, 57–8.

SMITH, W. M. (1965*b*). Visual recognition: Facilitation of seeing by hearing. *Psychonom. Sci.* **2**, 157–8.

SOLOMON, R. L. & HOWES, D. H. (1951). Word frequency, personal values, and visual duration thresholds. *Psychol. Rev.* **58**, 256–70.

SOLOMON, R. L. & POSTMAN, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.* **43**, 195–201.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Columbia University Press.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

WINNICK, W. A. & KRESSEL, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning and free recall as a function of abstractness–concreteness and word frequency. *J. exp. Psychol.* **70**, 163–8.

ZAJONC, R. B. & NIEUWENHUYSE, B. (1964). Relationship between word frequency and recognition: Perceptual process or response bias? *J. exp. Psychol.* **67**, 276–85.

(*Manuscript received* 15 *May* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, 413–419
Printed in Great Britain

EXPERIMENTAL STUDY OF THE EFFECT OF LETTER REVERSALS ON READING

By DOREEN ASSO and MARIA WYKE
Department of Neurosurgical Studies, Institute of Neurology, Queen Square, London

A study of the effect of letter reversals (mirror reversals) on the direction of reading words exposed tachistoscopically was carried out. The subjects were twenty experienced readers, and the test material consisted of words which could be read meaningfully in the right–left as well as in the left–right direction. The words were presented with and without letter reversals. The results show that, under these conditions, subjects read from right to left the words containing the reversed letters. The effect of word frequency, as well as the position within the word of reversed letters, has also been analysed. Some theoretical interpretations of the findings are discussed.

This paper reports an experiment designed to analyse the effect of letter reversals on the direction of reading.

Two main types of eye-movement have been described in subjects reading English (Carmichael & Dearborn, 1947). The first involves a series of short ocular movements from left to right along the line of print. The second consists of movements from right to left at the end of each line. In the first type of movement, two tendencies control the deflexions of the eyeballs: one tendency involves fixation on the beginning of the line of print and the other results in saccadic ocular jumps from left to right. When the English reader looks at a line of print, the tendency to start at the beginning of a line is considered to be dominant.

Left–right attentional scanning processes associated with the reading of printed material have also been demonstrated when letters or words are exposed tachistoscopically for periods of time brief enough to prevent eye-movements. Thus several experiments designed to compare the efficiency of recognition in the two halves of the binocular field of vision have shown that letters and words are identified more accurately in the left field, provided that the material is presented simultaneously in both fields. On the other hand, recognition is better in the right visual field when the letters or words are presented successively, i.e. first in one field then in the other (Bryden & Rainey, 1963; Heron, 1957; Mishkin & Forgays, 1952; Terrace, 1959). It has been suggested that the perceptual superiority of the left visual field during bilaterally simultaneous presentation of written material in the two fields of vision is related to reading experience (Bryden, 1960; Heron, 1957). Experienced readers will attend to the letters or words in the order in which they would normally read; that is to say, letters which would be fixated first under normal reading conditions will be scanned first.

The left–right direction in reading depends on the fact that the English language has direction incorporated in the structure of written material. The phonetic order of word formation is from left to right; also, the sequence of word meaning is from left to right, and furthermore a large number of English words have the etymological root which is the primary carrier of the meaning at the beginning of the word (Harcum & Dyer, 1962). However, one factor whose influence in this regard has not

been clarified is the significance of the orientation of the individual letters in the development of the left–right reading habit.

Only eleven of the letters in the English alphabet are symmetrical. It is therefore possible that the left–right orientation of the letters within a word might provide a cue for the direction in which a word should be scanned; and this possibility has been previously analysed experimentally. Thus Wolfe (1939) observed that children between the ages of 9 and 11 years (and with concordant reading ages) will read reversible words from right to left when such words were presented with letters reversed (using mirror reversals). A reversible word, being one that can be read meaningfully from left to right and from right to left, provides no conflicting opposition between meaning and direction. However, this experiment lacked control for set, for all the reversible words were presented with mirror reversals; and the results were therefore interpreted in terms of the acquisition of a 'set to read backwards'.

More recently, Harcum & Filion (1963) and Harcum & Smith (1963) have studied the effect of letter reversals on the recognition of eight-letter words. The results of these experiments were evaluated in terms of errors of letter- rather than of word-identification; and this fact, as well as some limitations of the experimental design, has made it difficult to understand the role of letter orientation in the left–right reading habit.

In the present paper an attempt is made to analyse the possible role of right–left reversal in the identification of reversible words.

The purpose of this study is to decide to what extent the direction of non-symmetrical letters influences the left–right reading habit of experienced readers.

METHOD

Subjects

The subjects were twenty English women students between nineteen and twenty-two years of age (mean age 20·2). Their visual acuity was within normal limits or had been corrected with spectacles. All were right-handed.

Test material

This consisted of 100 cards, each having three- or four-letter words on it. The words were printed with black ink in capital letters. Eighty-eight of these words were reversible; that is to say, they could be read meaningfully from left to right and from right to left (e.g. FLOW). Twelve of them were non-reversible, i.e. they could be read meaningfully only in the left–right direction (e.g. MAN). The reversible words were divided into two groups.

Group 1. Words with right–left reversal of all the non-symmetrical letters, i.e. mirror reversals. There were twenty-two words together with their corresponding twenty-two reversible forms. Thus the word ꟻꓶOW was also presented as WOꓶꟻ; the word ꓷƎW was also shown as WƎꓷ, and so forth.

Group 2. The same twenty-two words and their reversible forms as in group 1, but with the letters normally orientated, i.e. FLOW, WOLF; DEW, WED. Of the remaining twelve non-reversible words, six had normally orientated letters, the other six had letters with mirror reversals. These twelve words were included in order to avoid the formation of a set in the direction of reading.

Apparatus

The words were presented with a Dodge-type tachistoscope at the centre of the field of vision. The words subtended a horizontal angle of 2° at the nodal point of the eye. The experiments were conducted in a darkened room under conditions of prior adaptation. During exposure of the cards, the luminance at the screen surface was 1·9 ft.-lamberts.

Procedure

The 100 cards were presented to all subjects in a random order, subject to the restriction that no word should be followed or preceded by its reversible form; e.g. WED was never preceded or followed by the word DEW. Four different random orders were used for the twenty different subjects.

Each word was first exposed for 7·5 msec. If the subject failed to identify it, it was presented again at an increased exposure time. The time exposures ranged from 7·5 to 75 msec in the following steps: 7·5, 9·5, 12, 15, 20, 25, 30, 40, 50, 60, 75.

The subjects were instructed to keep both eyes open and, when told, to look at a black dot placed at the centre of the screen. They were asked to read aloud the word they saw immediately after each presentation.

Three trial cards were used. The words in these cards included non-reversible words, and there was no alteration in the orientation of the letters.

The subjects were not warned of the changes of letter orientation within some of the words, nor that some words could be read meaningfully from right to left. If the subjects commented upon the letter reversals they were told that they must attempt to read the word in spite of the changes in orientation. A response was considered correct when the subject identified the word accurately in either the left–right or right–left direction; the total number of words read in either direction was then recorded.

Results

For scoring purposes the exposure values in milliseconds were assigned to a position on an arithmetic scale varying from 5 to 15. Thus the shortest exposure time, 7·5, was placed in position 5 and the longest in position 15.

The mean exposure time for recognition of the 100 cards for the twenty subjects was 9·5. There was no significant difference in the mean exposure time for recognition between the words containing letter reversals and those with letters normally orientated.

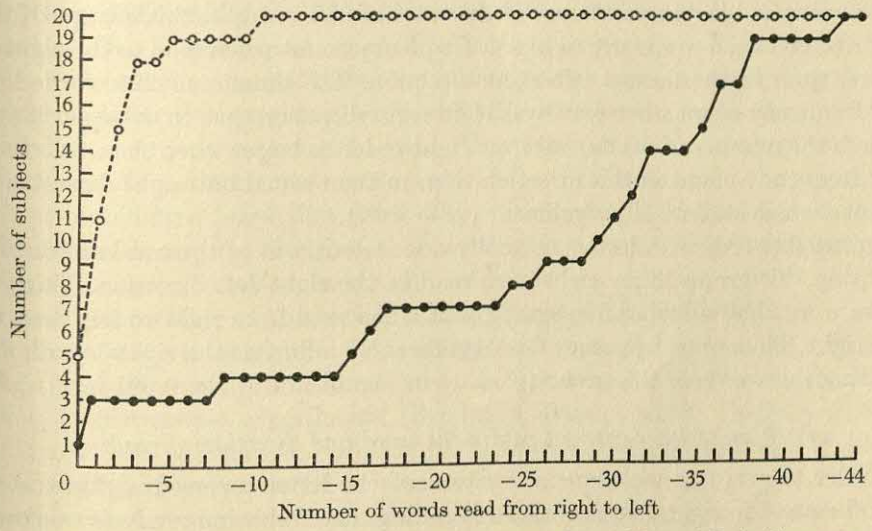


Fig. 1. Cumulative frequency distribution of words read from right to left in the group of words with reversed letters (●—●) and in the group with letters normally orientated (○ - - ○).

The cumulative frequency distribution of words read from right to left in the two groups of reversible words is shown graphically in Fig. 1. In group 1 (i.e. with letter reversals) only one subject read all the words in the normal left-to-right direction.

Half of the subjects (10) read thirty or more words from right to left, and one subject read all words (44) in the right–left direction. The median number of words read from right to left in this group is 29. In group 2 (i.e. with letters normally orientated) five subjects read all words in the normal direction (i.e. left to right); fourteen subjects read five words or less from right to left; one subject read ten of the forty-four words in the right-to-left direction. The median number of words read from right to left in this group is 0·8.

Word frequency and direction of reading

The words of each pair presented to the subjects vary in frequency of usage according to the direction in which they are read. Thus the words TON and NOT are relatively different in frequency of usage.

In order to evaluate whether word frequency was a factor in determining the right–left direction of reading, the following analysis was carried out.

On the basis of the Thorndike–Lorge count (Thorndike & Lorge, 1944), the twenty-two words and their reversible pairs were classed as 'more frequent' and 'less frequent'. A 'more frequent' word was a word with a higher frequency value when read in the normal left-to-right direction than when read from right to left. A 'less frequent' word was one with a lower frequency value when read in the left–right than in the right–left direction. Thus the word NOT was classed as 'more frequent' because its frequency in the Thorndike–Lorge count occurs at least 100 times per million, while the frequency of the reversible word TON is only 49 per million. On the same basis the word RAW was classed as 'less frequent', as it occurs less frequently in the word count than its reversible form WAR. Twenty-two words were classed as 'more frequent' and twenty-two as 'less frequent'.

In group 1 (words with letter reversals) the number of words read from right to left by the twenty subjects was 494 out of the total of 880 words in this group. Of these 494 words, 269 were words with a higher frequency count when read in the right–left direction than in the normal left–right direction. The remaining 225 words had a higher frequency count when read in the left–right direction than in the opposite way. Although the number of words read from right to left is larger when the words have a higher frequency count in this direction than in the normal left–right direction, the difference is not statistically significant ($\chi^2 = 3{\cdot}57$).

In group 2 (words with letters normally orientated), out of the total of 880 words comprising this group thirty-eight were read in the right–left direction. Thirty-one of these words had a higher frequency count when read from right to left than from left to right. Thus word frequency does significantly influence the right–left direction in reading the words in this group ($\chi^2 = 15{\cdot}01$ significant at the 0·001 level).

Position of letter reversal within the word and direction of reading

In order to evaluate whether the position of the letter reversed within the word will influence the right–left direction of reading, the following analysis was carried out:

The forty-four words with letter reversals were divided into three groups: (i) words with a letter reversed at the beginning of the word, e.g. ꟻOT, (total 15); (ii) words with a letter reversed at the end of the word, e.g. TAꟻ (total 15); (iii) words with letters reversed at both the beginning and the end of the word, e.g. ꟻIVꟻ (total 14).

Out of the 494 words read from right to left by the twenty subjects, 170 had a letter reversed at the beginning of the word, 143 ended with a letter reversal, and 181 had letter reversals at both ends. Although more words were read from right to left when the letter reversals appeared at both the beginning and the end of the word than when they appeared only at either the beginning or the end, the difference is not statistically significant ($\chi^2 = 3{\cdot}65$).

Discussion

This study shows that when experienced readers are presented (at brief exposure times) with words that can be read meaningfully in the right–left as well as in the left–right direction, they will read from right to left if the words contain reversals of the non-symmetrical letters. Furthermore, this alteration in the direction of reading is not significantly related to the frequency of the word containing the reversed letters. Thus an experienced reader, when confronted with a word containing reversed letters, reads it from right to left, in spite of the fact that the word read in this direction might be less frequent than that read in the normal left–right direction.

The results also show that the number of words read from right to left is not significantly associated with the position of the reversed letters within the word. Thus reading from right to left will occur with equal frequency, irrespective of whether words have the reversed letters at the beginning, at the end or at both the beginning and the end of the word.

These observations imply, then, that letter orientation does provide an important cue for the direction in which the word should be read. Pertinent in this respect are the observations of Gibson *et al.* (1962) on the development, between the ages of 4 and 8 years, of the ability to discriminate visually a set of letter-like forms as well as real letters. One of the aspects studied was the discrimination of letter-like forms from their right–left reversals. It was found that the errors in discriminating reversals were higher in the younger group, but had declined almost to zero by the eighth year. This finding might be expected because directional orientation is not critical for identification of the different letter-like forms. It seems that between the seventh and eighth years children learn that orientation of letters is a factor relevant to their discrimination. It may be, therefore, that learning the correct orientation of letters and learning to read words are strongly connected.

In the interpretation of the results of the present study, it is relevant to inquire whether or not the alteration in the reading direction in the presence of words containing reversals involves an alteration in the normal attentional *scanning* habits. Previous tachistoscopic experiments (Bryden & Rainey, 1963; Harcum & Filion, 1963; Heron, 1957; Mishkin & Forgays, 1952; Terrace, 1959) with printed alphabetic material regard the attentional processes involved in letter recognition and those involved in word recognition as somewhat similar. It is reasonable to assume, however, that the recognition of a series of unrelated letters and the identification of a word involve two different visual attentional processes. In studies of letter recognition, the subjects are directed to attend to each letter in turn; and in such circumstances, the order in which the subject will attend to the letters depends upon certain features of the experimental situation (Gibson, 1963).

On the other hand, when a subject reads a word he does not attend to each letter in turn. Studies of the visual attentional processes during normal reading clearly indicate that the eyes do not fixate every letter (Carmichael & Dearborn, 1947). Instead, the experienced reader perceives certain characteristics of the line of print that serve as cues to the words contained therein; in short, it is the whole word rather than its components that is perceived—and this will apply *a fortiori* in the case of tachistoscopic presentation of words containing relatively few letters. In the light of these data, it is possible that by reversing the letters within a word what is altered is the perception of the distinctive appearance of the word, rather than the direction of the attentional scanning habits. In other words, reversed letters may provide certain cues that make it possible to identify a complete word in a right–left orientation.

This interpretation is reinforced by the analysis of the effect of the position of the reversed letters within the words. In words ending with a reversed letter, this final letter would constitute the first visual cue in a right–left scanning sequence; and if the direction of the scanning processes were altered by the presentation of the words containing reversed letters, it would be expected that a larger number of words with a terminal reversed letter would be read from right to left than when the reversed letter was at the beginning of the word. However, the results of the present study show that there is no significant association between the position of the reversed letter within a word and the number of times a word was read from right to left.

The observation that words are perceived as entities also explains why a few words in the present experiment were read from right to left, in spite of having the letters normally orientated. This arose in the few instances in which the frequency value of the word read from right to left was considerably higher than that read in the normal left–right direction: for familiarity provided an important cue, and the subject perceived the more familiar of the two words.

The distinction between the attentional processes involved in individual letter recognition and those associated with reading of a complete word is important in the evaluation of the 'strength' of the left–right attentional reading habits—particularly as the 'strength' of such habit has been used to explain the relative superiority of the left visual field of vision during the simultaneous presentation of stimuli in both fields (Bryden, 1960; Heron, 1957). It is also clear that future studies of the role of letter orientation in determining the left–right scanning habits associated with reading should take into account the limitations imposed by the use of short familiar words, long series of unconnected letters, or words which have no meaning when read from right to left.

We should like to thank Mr V. Logue for facilities provided, the students of the West End Hospital School of Speech Therapy who acted as subjects, and the Department of Psychological Medicine, Maida Vale Hospital, for the loan of apparatus. We are also grateful to Mr A. H. Prickett for the photographic work.

The work was supported by grants from the Institute of Neurology, Queen Square, and the Medical Research Council.

References

Bryden, M. P. (1960). Tachistoscopic recognition of non-alphabetic material. *Can. J. Psychol.* **14**, 78–86.

Bryden, M. P. & Rainey, C. R. (1963). Left–right differences in tachistoscopic recognition. *J. exp. Psychol.* **66**, 568–71.

Carmichael, L. & Dearborn, W. H. (1947). *Reading and Visual Fatigue.* Boston: Houghton Mifflin.

Gibson, E. J. (1963). Perceptual learning. *Ann. Rev. Psychol.* **14**, 29–56.

Gibson, E. J., Gibson, J. J., Pick, A. D. & Osser, H. (1962). A developmental study of the discrimination of letter-like forms. *J. comp. physiol. Psychol.* **55**, 897–906.

Harcum, E. R. & Dyer, D. W. (1962). Monocular and binocular reproduction of binary stimuli appearing right and left of fixation. *Am. J. Psychol.* **75**, 56–65.

Harcum, E. R. & Filion, R. D. (1963). Effects of stimulus reversals on lateral dominance in word recognition. *Percept. mot. Skills* **17**, 779–94.

Harcum, E. R. & Smith, N. F. (1963). Effect of pre-known stimulus-reversals on apparent cerebral dominance in word association. *Percept. mot. Skills* **17**, 799–810.

Heron, W. (1957). Perception as a function of retinal locus and attention. *Am. J. Psychol.* **70**, 38–48.

Mishkin, M. & Forgays, D. G. (1952). Word recognition as a function of retinal locus. *J. exp. Psychol.* **43**, 43–8.

Terrace, H. (1959). The effects of retinal locus and attention on the perception of words. *J. exp. Psychol.* **58**, 382–5.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words.* New York: Columbia University Press.

Wolfe, L. S. (1939). An experimental study of reversals in reading. *Am. J. Psychol.* **52**, 533–61.

(*Manuscript received* 5 *April* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 421–433
Printed in Great Britain

RECENT DEVELOPMENTS IN SHORT-TERM MEMORY

By BENNET B. MURDOCK, Jr.
University of Toronto

This paper reviews some recent developments in the area of short-term memory. The methodological innovations include the distractor technique and the probe technique. The main empirical phenomena discussed are recency effects, input and output interference, the length–difficulty relationship, inter-item interference, proactive interference effects, the distribution of practice interaction, associative symmetry, modality effect, acoustic confusions, and transpositions. A 'modal model' is presented which attempts to synthesize some recent theoretical conceptions; the components include sensory, short-term and long-term stores with three different forgetting mechanisms (decay, displacement and interference, respectively).

Recent developments in the field of short-term memory (STM) have provided a 'new look' at many of the old problems in the field of human verbal learning and retention. There are new techniques of experimental investigation, new empirical phenomena and new theoretical models. In general, quite a different picture of human memory is emerging, and the concepts of information processing are coming to play an increasingly important role. This paper is a review of some of the more important recent developments in the field and is organized around the three topics mentioned above: techniques, phenomena and models. For earlier articles on these topics see Broadbent (1965), Brown (1964), Peterson (1966*b*), Posner (1963, 1967) and Postman (1964).

Techniques

With discrete trials there are two main techniques for the study of STM: the distractor technique and the probe technique. The former is characterized by the single presentation of a sub-span item, followed by some irrelevant activity which is designed to prevent rehearsal. The latter is characterized by the presentation of a *list* of items followed by a post-stimulus cue for one of the items. The distractor technique was first introduced by John Brown in 1958 and the Petersons in 1959; Brown (1958) varied the *amount* of material to be remembered (one to four consonant digrams) with a fixed retention interval (five two-digit numbers to be read aloud), whereas Peterson & Peterson (1959) presented a single CCC trigram and varied the *duration* of the irrelevant activity (counting backwards by three's or four's over a range of 3–18 sec).

The probe technique was used by Anderson (1960) with digit strings, where lists of twelve digits were presented in blocks of four and the subjects were variously cued to recall the first block, the second block or the third block. Also Sperling (1960) used tachistoscopic arrays of letters with the same sort of post-stimulus cueing. As both authors indicated, they were not the first to use partial report; what did uniquely characterize their studies was use of a *sampling* technique to cue recall. Murdock (1961*b*) used a probe technique for paired associates; Waugh & Norman (1965) tested the recall of single digits from a string of digits and, of course, gave us the name for the technique.

It has been argued (e.g. Keppel, 1965) that there is somehow a basic difference between the distractor technique and the probe technique, since in the one case the retention interval is filled with categorically different material which the subject

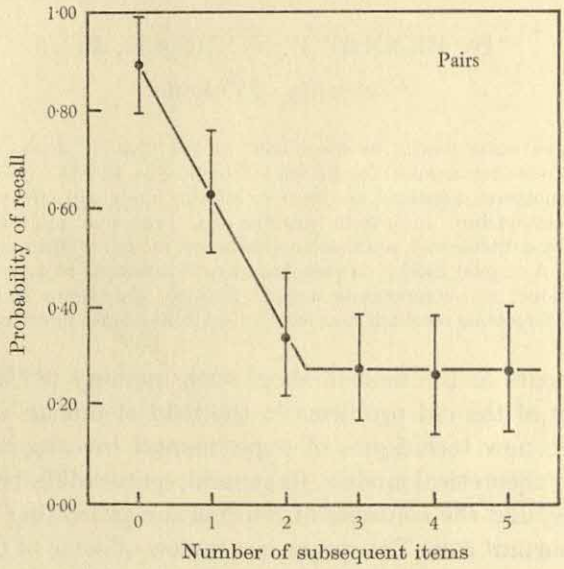


Fig. 1. STM curve from probe technique (data from Murdock, 1967*b*).

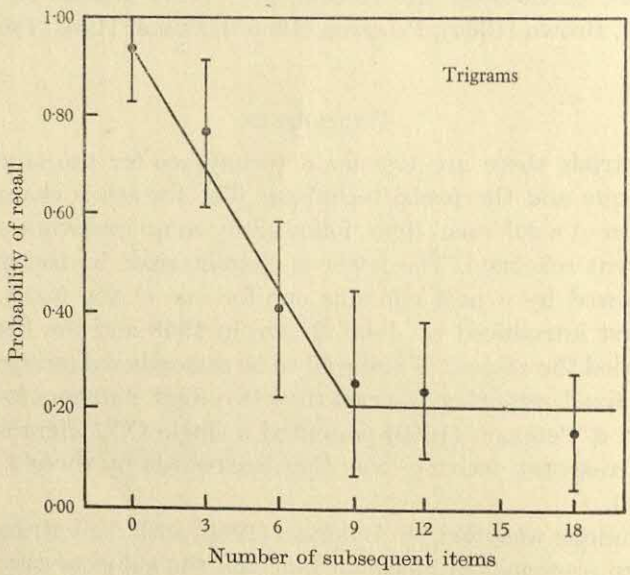


Fig. 2. STM curve from distractor technique (data from Murdock, 1961*a*).

need not learn, whereas in the other case the retention interval is filled with the same type of material which the subject must try to learn. Contrary evidence is shown in Figs. 1 and 2, where Fig. 1 shows retention data from a probe technique with paired associates and Fig. 2 shows retention data from a distractor task with CCC trigrams. In both cases probability of recall is plotted as a function of number of subsequent

items (which are other pairs in one case and three-digit numbers in the other case): in both cases the vertical lines indicate the magnitude of the S.D.s across subjects. The two straight lines may not be the 'true' retention function, but there is little doubt that the data are remarkably similar under these two logically different procedures. It is suggested then that the distractor and the probe techniques give essentially equivalent results; for further discussion of the point see Murdock (1967*b*).

Broadbent (1965) has distinguished not only probe and distractor techniques but also, for the latter, predictable and unpredictable interpolation. Also, there are other techniques for the study of STM: the classical memory span (Crannell & Parrish, 1957), running memory span (Nixon, 1946; Pollack, Johnson & Knaff, 1959), multichannel presentation (Broadbent, 1958; Moray, Bates & Barnett, 1965), and continuous memory tasks (Lloyd, Reid & Feallock, 1960; Shepard & Teghtsoonian, 1961; Yntema & Mueser, 1960). The continuous memory tasks for some purposes may be very useful. However, they do suffer from the fact that input and output effects are generally confounded (i.e. the retention interval is filled with the presentation of new items and/or recall of old items) and criterion changes may necessitate an analysis based on signal-detection theory (Donaldson & Murdock, 1968). Distractor and probe techniques focus on the retention of single items in a discrete-trials procedure, and results presented here will be drawn largely from studies using these techniques.

Empirical phenomena

In this section ten of the characteristic phenomena of STM are described; in general they are large and dramatic effects, and clearly reproducible in different laboratories. Essentially all that will be possible in the available space is briefly to describe each one and indicate what its general significance is. The effects to be described are: recency effects, input and output interference, length–difficulty relationship, inter-item interference, proactive interference effects, distribution of practice interaction, associative symmetry, modality effects, acoustic confusions, and transpositions.

Recency effects

The most recently presented items will be the best remembered, and the more remote the item the poorer the retention. This effect has already been indicated in Figs. 1 and 2, if it is remembered that low numerical values of 'number of subsequent items' correspond to recently presented items. The drop-off is very precipitous or the recency effect is very pronounced. Further indication of the generality of the effect is illustrated in Fig. 3, which shows serial position curves for free recall for lists of varying length and different presentation rates (see Murdock, 1962, for further details). Again the most recently presented items are consistently the best retained; the methods of presenting the data may change but the nature of the effect is everywhere the same.

Input and output interference

There is a reduction in recall regardless of whether a given item is followed by the presentation of an additional item (input interference) or of a recall test for a prior item (output interference). The term input interference is merely another way of talking about the recency effects noted above; clear evidence for output interference may be

found elsewhere (Murdock, 1963*b*; Tulving & Arbuckle, 1963). The magnitude of the effects is comparable; Waugh & Norman (1965) assume them to be the same, though in a later study Tulving & Arbuckle (1966) report input interference to be more disruptive than output interference.

A striking demonstration is reported by Conrad (1958), who found that even a single redundant digit can, under appropriate conditions, markedly reduce performance. His subjects were skilled telephone operators who merely had to dial the digit 0 before recalling the string; in this and in a later experiment (Conrad, 1960) performance fell off by perhaps 50 % as a consequence. Thus STM is extremely transitory, and it almost would seem as if *anything* interpolated between input and output will depress performance.

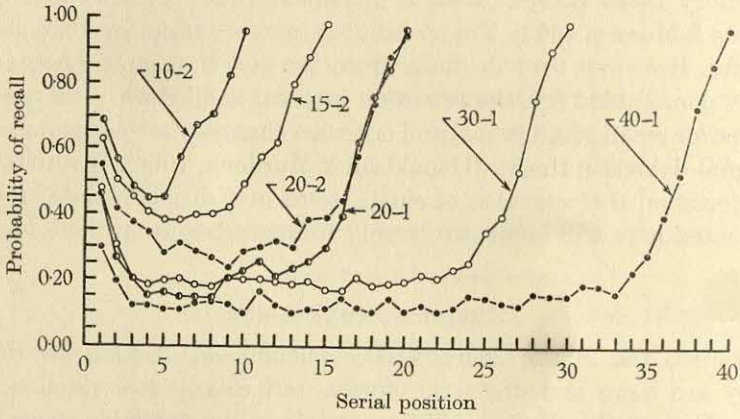


Fig. 3. Serial position curves from free recall (data from Murdock, 1962).

Length–difficulty relationship

The retention curve for STM approaches an asymptote greater than zero. Therefore the total number of items recalled after a single presentation should be a linear function of list length. This argument is detailed more fully elsewhere (Murdock, 1963*a*); the point is that since any 'early' list item has a fixed recall probability total *number* of items recalled when all are tested should be linear when plotted against number of items in the list. Evidence from free recall (Murdock, 1960) suggests that the effect holds over an extensive range.

Data from paired-associate learning (Murdock, 1967*c*) are shown in Fig. 4. These data are based on two willing volunteers who alternated as subject and experimenter; they learned paired-associate lists ranging in length from 9–100 pairs (common English words paired at random) with a training and testing method. Here only first-trial recall is presented; as is evident, number of pairs recalled after a single presentation did increase with list length and a linear function is not unreasonable. Further, since the presentation time was 2 sec/pair during the input phase, the predicted slope of the regression line would be 0·25; a least-squares determination gave a numerical value of 0·267. Apparently one really can extrapolate quite successfully from retention intervals measured in seconds to retention intervals measured in minutes.

Intra-item interference

The more items there are in the sub-span unit the poorer will be the retention. A single word or a single letter is forgotten slowly, three words or three letters more rapidly, and five more rapidly still (Melton, 1963; Murdock, 1961*a*). These findings suggest that a CCC trigram, for instance, is not a single item. They call attention to the units in which the stimulus material is specified, and emphasize again the importance of the coding problem in STM (Miller, 1956). For a fuller discussion of this point see Melton (1963).

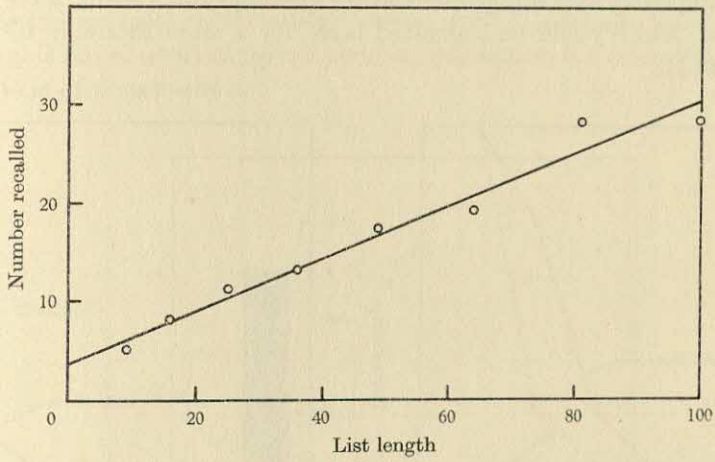


Fig. 4. Number of pairs recalled as a function of list length (data from Murdock, 1967*c*).

Proactive interference effects

Prior items interfere with the retention of subsequent items. Underwood (1957) has shown that, in long-term memory, the naive subject remembers 75 % of the material over a 24 hr retention interval, whereas the experienced subject remembers only 25 % of the material. A comparable effect in STM was demonstrated by Keppel & Underwood (1962); on the very first trial subjects showed almost no forgetting in a distractor task but by three or four trials the retention curves were comparable to those of highly practised subjects. Wickens, Born & Allen (1963) even find a 'release from PI'; shifting from one class of stimulus material (e.g. trigrams) to a different class of material (e.g. digits) appears to wipe the slate clean.

In looking for a comparable effect with paired associates we found instead a see-saw effect. That is, with practice, retention of the first pair does decrease, but there is a compensatory increase in retention of the penultimate pair (Murdock, 1964). Why the discrepancy between trigrams and pairs? One possibility is that the first pair in a list is analogous to the trigram, while later pairs are analogous to the interpolated activity. If so, the paired-associate data would predict a compensatory effect for retention of the interpolated material in the distractor technique, but this retention is never measured. However, under appropriate conditions it can be, and in one experiment (Murdock, 1966*b*) the results were clearly confirmatory. That is, with practice, retention of the trigram goes down but retention of the interpolated material

does go up and, percentage-wise, the effects balance nicely. So 'proactive interference' is only part of the story; there is a change from primacy to recency with practice, and in a constant amount of time the subject can remember a constant amount of material.

Distribution of practice interaction

This is 'the Peterson paradox': with two presentations of an item, massed practice is better at long retention intervals. The 'paradox' is this: since forgetting occurs between the first and the second presentation of an item, the interaction means that to maximize long-term retention after the second presentation, one wants to *forget* the item as completely as possible after the first presentation. Clearly there are various interpretations which could be suggested here; for a more thorough discussion see Peterson (1966*b*).

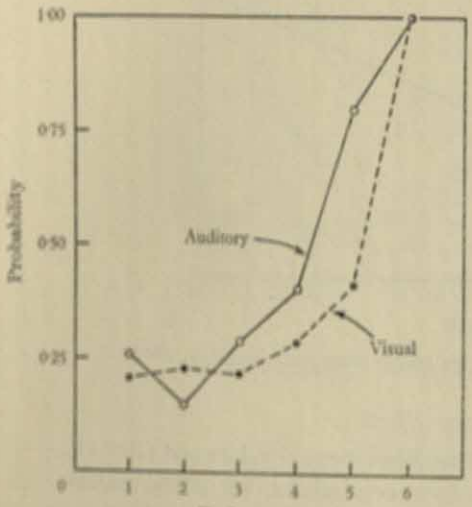


Fig. 5

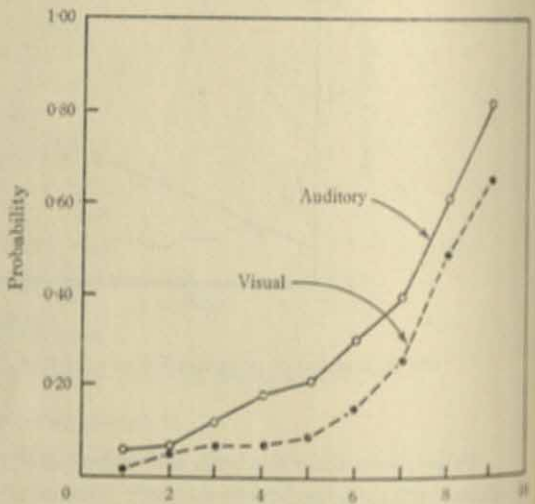


Fig. 6

Fig. 5. Recall probability for auditory and visual pairs (data from Murdock, 1967*a*).
Fig. 6. Recall probability for ten-word serial presentation (data from Murdock, 1967*a*).

Associative symmetry

Following a single presentation of an A–B pair, backward recall (i.e. recall of A given B as the probe) is as good as forward recall (i.e. recall of B given A as the probe). The effect is best demonstrated with dichotic presentation (one word of the pair to the left ear, the other word to the right ear); in this way one can be sure that the input is in fact simultaneous. The effect holds regardless of whether one defines 'forwards' and 'backwards' in terms of the order of verbalization by the subject himself or in terms of normative data on linguistic preferences obtained from an independent sample of subjects (Murdock, 1966*a*). The theoretical issues involved in associative symmetry have been extensively discussed by Asch & Ebenholtz (1962), and the empirical data referred to above cast grave doubt on the utility of a classical conditioning model applied to individual associations in STM.

Modality effects

With verbal material, auditory presentation results in superior retention to visual presentation. With paired associates the advantage seems to be localized near the end of the list but with serial presentation the effect is more extensive, being demonstrated over all probe positions (Murdock, 1967*a*). Illustrative data are shown in Figs. 5 and 6; it should be remembered that conditions of presentation were designed to be as comparable as possible. The direction of the effect may reflect primarily the fact that, in STM, discrimination on the basis of temporal relationships is required; had the distribution been spatial the situation might have reversed. These findings demonstrate the importance of 'sensory stores' in STM, and also show quite clearly that retrieval is not ahistorical; given veridical perception the retention still depends upon the mode of presentation.

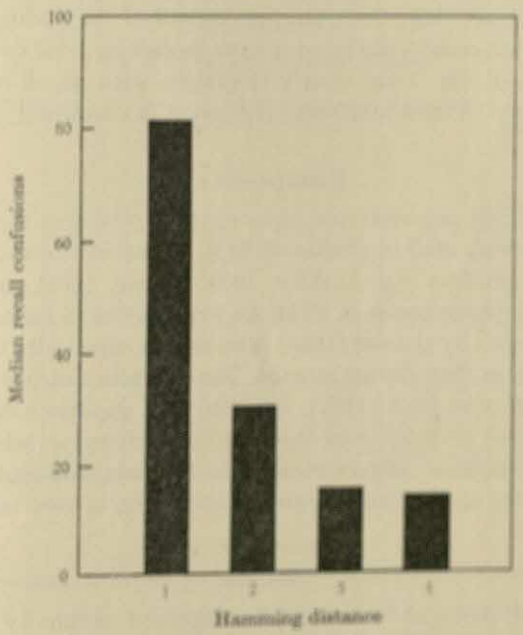


Fig. 7. Recall confusions as a function of Hamming distance (data from Conrad, 1964).

Acoustic confusions

Items which are acoustically similar (i.e. are confused on a listening test) are harder to remember in STM even when presentation is visual. Evidence for this effect comes from Conrad (1964) and Wickelgren (1965*a*); apparently the magnitude is considerable. To date the confusion effect has been found both for consonants and for vowels, and it may be reducible to a 'distinctive features' analysis (Wickelgren, 1965*b*, 1966).

As a demonstration, we took the Conrad data on single substitution errors (Conrad, 1964) and, for each of the two subsets separately (BCPTV; FMNSX) applied a Miller-Nicely (1955) distinctive-features analysis to determine the number of features on

which each pair of consonants differed (e.g. *B* and *D* differ only in place of articulation but *P* and *D* differ both in voicing and place of articulation). The average (median) number of recall confusions for the two subsets combined is shown in Fig. 7 as a function of number of distinctive features in which they differ; the label 'Hamming distance' is a term from coding theory (Abramson, 1963) to specify similarity between pairs of binary-coded numbers. As can be seen, the most common recall confusion is between consonants which differ in only one distinctive feature, and confusions decrease monotonically with dissimilarity.

The significance of the acoustic confusion effect is threefold: (*a*) Originally it was held (e.g. Broadbent, 1963) that similarity was generally an ineffective variable in STM, and this belief was used as an argument by decay theorists. We now know that it is wrong; previous studies were simply looking at the wrong dimensions of similarity. (*b*) The acoustic (or articulatory) effect strongly suggests that representation in STM is through a speech or speech-motor code. Whether it is the acoustic or the articulatory feedback is not yet clear, but some component of the auditory process clearly predominates. (*c*) Forgetting goes on at a more molecular level even than the single item, and the data of Fig. 7 are clearly consistent with an all-or-none concept of forgetting at the level of item attributes (Bregman & Chambers, 1966; Wickelgren, 1965*b*).

Transpositions

With serially ordered presentations, place changes (transpositions) occur over the brief retention intervals used in studies of STM. Transpositions are a manifestation of the serial-order problem (e.g. Lashley, 1951; Milner, 1961) and they are a consistent but puzzling phenomenon in STM. An explanation in terms of item substitution has been proposed by Conrad (1965), who argues essentially that transpositions occur at output rather than during storage. The opposite conclusion is suggested by data of Murdock & vom Saal (1967), who found a positive relationship between accuracy of recall and probability of transpositions. However, what does seem very clear is that transpositions are rare events which, given sufficient data, are regular and reproducible; the mechanism responsible, however, is very far from clear.

MODELS

There has recently come to be a surprising degree of unanimity among the model-builders in the area of STM. In what follows we shall try to present a 'modal model'; a synthesis of the various views recently put forth. Naturally this synthesis will do violence to some, but given the consensus the distortions are hopefully not too great. The ideas are particularly those of Broadbent (1958), Sperling (1963), Bower (1964, 1967), Waugh & Norman (1965), Atkinson & Shiffrin (1965) and Peterson (1966*a*).

The general model is shown in Fig. 8; items upon presentation enter the sensory store (SS); modality effects operate here, and multichannel inputs are possible. If the subject is 'attending' and the processing demands not too great, items pass through a single channel and enter primary memory (PM). PM has a fixed contents; it can hold only so many items. While items reside in PM they may also be copied or transferred into secondary memory (SM), and 'rehearsal' is a critical variable here. PM is also called a memory buffer or a short-term store; SM is a longer-term store

with persistence (at least in theory) of the order of minutes or perhaps hours. The upper bound has been neither specified nor investigated.

Three different 'forgetting' mechanisms may be suggested: *decay* (in the matter of milliseconds) from SS, *displacement* from PM and *interference* from SM. The displacement notion (see Atkinson & Shiffrin, 1965) is that of a push-down buffer which can hold exactly r items; when overflow occurs an item is knocked out, though not necessarily in strict order of arrival. Transfer into SM is by means of a limited capacity mech-

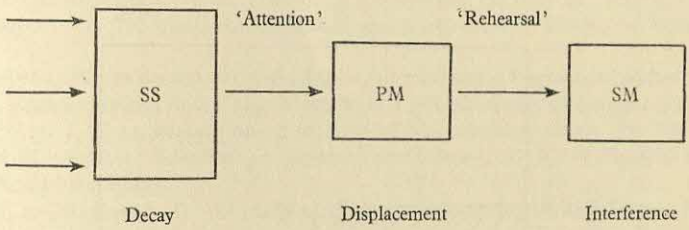


Fig. 8. Composite model for STM.

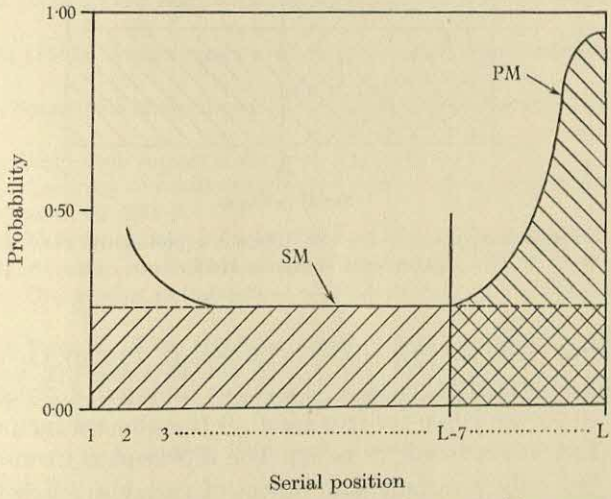


Fig. 9. Primary memory, secondary memory, and the serial position curve of free recall.

anism which operates in real time; SM is rate sensitive while PM is not. The exact nature of the interference in SM is not completely clear; it may be interference in the classical sense with confusions at the time of recall, proactive inhibition effects, and/or intra-item interference effects. In this connexion it should be noted that Tulving (1968) has suggested that the distinction should not be between dual storage systems but between different retrieval processes, and the limitation in capacity is in terms of retrieval capacity.

Items can reside in PM, SM, or both; the two storage systems are not mutually exclusive. This point is illustrated in Fig. 9, which tries to reconcile theory and data. The curve is an idealized serial position curve from free recall with list length L; PM includes (probabilistically) the last eight serial positions (with $r = 4$) and is cross-hatched down to the right. SM is represented by the asymptote that is so charac-

teristic of these curves (e.g. Fig. 3) and is crossed-hatched down to the left. The area of overlap indicates the dual representation. An experimental separation of PM and SM in free recall has recently been reported by Glanzer & Cunitz (1966). The primacy effect is generally neglected, but could be explicable in terms of longer residence in PM for those items entering an empty buffer. Finally, the rate-sensitive nature of SM is illustrated in Fig. 10; lists of 15–2 and 30–1 each have presentation times of $\frac{1}{2}$ min. The areas of the two rectangles should be equal; the ordinates are actually *data* though selected from three pair-wise comparisons (Murdock, 1962).

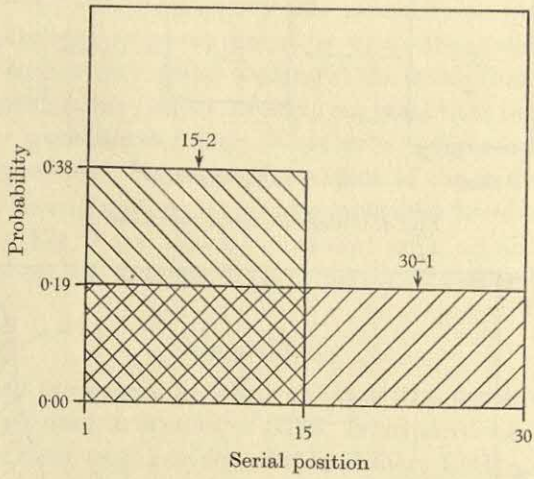


Fig. 10. Secondary memory for two different experimental conditions (data from Murdock, 1962).

We know from free recall that $R_1 = kt + m$ (Murdock, 1961*a*), i.e. number of items recalled after a single presentation (R_1), is a linear function of presentation time (t) with a slope of k (0·06 words/sec for free recall, 0·12 pairs/sec for paired associates) and intercept m (four to six words or pairs). The intercept is memory span, PM, or echo-box, and is generally invariant with a host of variables (presentation rate, list length, etc.). The rate constant, however, attests to the temporal dependence of SM, and in a general sense is a measure of information-processing ability of the subject. And in fact the numerical values are not grossly different from those conventionally found in studies of rate-of-gain and choice reaction time (Welford, 1960); as Leonard (1961) has suggested, coding compatibility may underlie both.

In conclusion, it should be stressed that this modal model is a composite of many different views which are better read in the original. It is not presented as the 'truth', and indeed the writer believes that some of it is actually wrong. While a separation of short-term and long-term stores is currently popular, Melton (1963) has argued effectively against such a dual storage model. However, at a general level a dual model is certainly consistent with a large amount of data, and it greatly helps to consolidate and integrate the disparate experimental findings which have so recently come to the scene. At the very least it gives us the illusion that we are finally talking about the processes and mechanisms which really underlie short-term memory.

This paper has been based on an invited address to Division 3 of the American Psychological Association, New York, 5 September 1966. Preparation of this paper was supported by Grant APA 146 from the National Research Council of Canada and by Grant GB 4545 from the National Science Foundation.

References

Abramson, N. (1963). *Information Theory and Coding*. New York: McGraw-Hill.

Anderson, N. S. (1960). Poststimulus cuing in immediate memory. *J. exp. Psychol.* **60**, 216–21.

Asch, S. E. & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proc. Am. phil. Soc.* **106**, 135–63.

Atkinson, R. C. & Shiffrin, R. M. (1965). Mathematical models for memory and learning. (Technical Report No. 79, Institute for Mathematical Studies in the Social Sciences, Stanford University.)

Bower, G. H. (1964). Notes on a descriptive theory of memory. (Paper read at Second Conference on Learning, Remembering, and Forgetting, New York Academy of Sciences, Princeton, N.J.)

Bower, G. H. (1967). A multicomponent theory of the memory trace. In *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed. K. W. Spence & J. T. Spence. New York: Academic Press.

Bregman, A. S. & Chambers, D. W. (1966). All-or-none learning of attributes. *J. exp. Psychol.* **71**, 785–93.

Broadbent, D. E. (1958). *Perception and Communication*. New York: Pergamon.

Broadbent, D. E. (1963). Flow of information within the organism. *J. verb. Learn. verb. Behav.* **2**, 34–9.

Broadbent, D. E. (1965). Techniques in the study of short term memory. *Acta psychol.* **24**, 220–33.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Q. J. exp. Psychol.* **10**, 12–21.

Brown, J. (1964). Short-term memory. *Br. med. Bull.* **20**, 8–11.

Conrad, R. (1958). Accuracy of recall using keyset and telephone dial, and the effect of a prefix digit. *J. appl. Psychol.* **42**, 285–8.

Conrad, R. (1960). Very brief delay of immediate recall. *Q. J. exp. Psychol.* **12**, 45–7.

Conrad, R. (1964). Acoustic confusions in immediate memory. *Br. J. Psychol.* **55**, 75–84.

Conrad, R. (1965). Order error in immediate recall of sequences. *J. verb. Learn. verb. Behav.* **4**, 161–9.

Crannell, C. W. & Parrish, J. M. (1957). A comparison of immediate memory span for digits, letters, and words. *J. Psychol.* **44**, 319–27.

Donaldson, W. & Murdock, B. B., Jr. (1968). Criterion changes in continuous recognition memory. *J. exp. Psychol.* (in press).

Glanzer, M. & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.* **5**, 351–60.

Keppel, G. (1965). Problems of method in the study of short-term memory. *Psychol. Bull.* **63**, 1–13.

Keppel, G. & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.* **1**, 153–61.

Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior*, ed. L. A. Jeffress. New York: Wiley.

Leonard, J. A. (1961). Choice reaction time experiments and information theory. In *Information Theory: Fourth London Symposium*, ed. C. Cherry, pp. 137–46. London: Butterworths.

Lloyd, K. E., Reid, L. S. & Feallock, J. B. (1960). Short-term retention as a function of the average number of items presented. *J. exp. Psychol.* **60**, 201–7.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.* **2**, 1–21.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–96.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. acoust. Soc. Am.* **27**, 338–52.

Milner, P. M. (1961). A neural mechanism for the immediate recall of sequences. *Kybernetik* **1**, 76–81.

MORAY, N., BATES, A. & BARNETT, T. (1965). Experiments on the four-eared man. *J. acoust. Soc. Am.* **38**, 196–201.

NIXON, S. R. (1946). Some experiments on immediate memory. (Applied Psychology Research Unit, A.P.U. 39.)

MURDOCK, B. B., JR. (1960). The immediate retention of unrelated words. *J. exp. Psychol.* **60**, 222–34.

MURDOCK, B. B., JR. (1961*a*). The retention of individual items. *J. exp. Psychol.* **62**, 618–25.

MURDOCK, B. B., JR. (1961*b*). Short-term retention of single paired associates. *Psychol. Rep.* **8**, 280.

MURDOCK, B. B., JR. (1962). The serial position effect of free recall. *J. exp. Psychol.* **64**, 482–8.

MURDOCK, B. B., JR. (1963*a*). Short-term memory and paired-associate learning. *J. verb. Learn. verb. Behav.* **2**, 320–8.

MURDOCK, B. B., JR. (1963*b*). Interpolated recall in short-term memory. *J. exp. Psychol.* **66**, 525–32.

MURDOCK, B. B., JR. (1964). Proactive inhibition in short-term memory. *J. exp. Psychol.* **68**, 184–9.

MURDOCK, B. B., JR. (1966*a*). Forward and backward associations in paired associates. *J. exp. Psychol.* **71**, 732–7.

MURDOCK, B. B., JR. (1966*b*). Measurement of retention of interpolated activity in short-term memory. *J. verb. Learn. verb. Behav.* **5**, 469–72.

MURDOCK, B. B., JR. (1967*a*). Auditory and visual stores in short-term memory. *Acta psychol.* **27**, 316–24.

MURDOCK, B. B., JR. (1967*b*). Distractor and probe techniques in short-term memory. *Can. J. Psychol.* **21**, 25–36.

MURDOCK, B. B., JR. (1967*c*). A fixed-point model for short-term memory. *J. math. Psychol.* **4**, 501–6.

MURDOCK, B. B., JR. & VOM SAAL, W. (1967). Transpositions in short-term memory. *J. exp. Psychol.* **74**, 137–43.

PETERSON, L. R. (1966*a*). Short-term verbal memory and learning. *Psychol. Rev.* **73**, 193–207.

PETERSON, L. R. (1966*b*). Short-term memory. *Scient. Am.* **215** (1), 90–5.

PETERSON, L. R. & PETERSON, M. J. (1959). Short-term retention of individual verbal items. *J. exp. Psychol.* **58**, 193–8.

POLLACK, I., JOHNSON, L. B. & KNAFF, P. R. (1959). Running memory span. *J. exp. Psychol.* **57**, 137–46.

POSNER, M. I. (1963). Immediate memory in sequential tasks. *Psychol. Bull.* **60**, 333–49.

POSNER, M. I. (1967). Short-term memory and information processing. *Acta psychol.* **27**, 267–84.

POSTMAN, L. (1964). Short-term memory and incidental learning. In *Categories of Human Learning*, ed. A. W. Melton. New York: Academic Press.

SHEPARD, R. N. & TEGHTSOONIAN, M. (1961). Retention of information under conditions approaching a steady state. *J. exp. Psychol.* **62**, 302–9.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* **74** (11). (Whole no. 498.)

SPERLING, G. E. (1963). A model for visual memory tasks. *Hum. Factors* **5**, 19–31.

TULVING, E. (1968). Free recall. In *Verbal Behavior and General Behavior Theory*, ed. T. R. Dixon & D. L. Horton. Englewood Cliffs, N.J.: Prentice-Hall.

TULVING, E. & ARBUCKLE, T. Y. (1963). Sources of intratrial interference in immediate recall of paired associates. *J. verb. Learn. verb. Behav.* **1**, 321–34.

TULVING, E. & ARBUCKLE, T. Y. (1966). Input and output interference in short-term associative memory. *J. exp. Psychol.* **72**, 145–50.

UNDERWOOD, B. J. (1957). Interference and forgetting. *Psychol. Rev.* **64**, 49–60.

WAUGH, N. C. & NORMAN, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

WELFORD, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics* **3**, 189–230.

WICKELGREN, W. A. (1965*a*). Acoustic similarity and retroactive interference in short-term memory. *J. verb. Learn. verb. Behav.* **4**, 53–61.

WICKELGREN, W. A. (1965*b*). Distinctive features and errors in short-term memory for English vowels. *J. acoust. Soc. Am.* **38**, 583–8.

WICKELGREN, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *J. acoust. Soc. Am.* **39**, 388–98.

WICKENS, D. D., BORN, D. G. & ALLEN, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *J. verb. Learn. verb. Behav.* **2**, 440–5.

YNTEMA, D. B. & MUESER, G. A. (1960). Remembering the present state of a number of variables. *J. exp. Psychol.* **60**, 18–22.

(*Manuscript received* 15 *May* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, 435–443
Printed in Great Britain

THE ROLE OF LANGUAGE AND EXPERIENCE ON THE USE OF LOGICAL SYMBOLS

By JAMES YOUNISS and HANS G. FURTH
Center for Research in Thinking and Language, Catholic University of America

Twenty-four deaf, twenty-four hearing culturally deprived (American Indian), and twenty-four middle-class control young adolescents were trained in five daily sessions to identify picture-instances as true or false denotations of symbolic statements of eight logical concepts. Training was non-verbal in that subjects had to construct instances for statements of concepts or write statements for given instances. On the fifth day subjects were administered a paper and pencil test in which they had to identify newly introduced instances as true or false cases of concept statements. Deaf and Indian children manifested a similar performance pattern across concepts which differed notably from that of the control middle-class subjects. With concepts which generated a broad class of logically equivalent but physically distinct instances and with that aspect of a concept which included a broad class of instances, control subjects were superior to deaf and to Indian subjects. In distinction, minimal differences among groups were observed in identification of an instance which came from a class of one or a narrow class. It was concluded that deaf and Indian subjects at the age-level sampled tended to use symbols primarily to depict attributes literally while the controls were better able to apply symbols as they represent classes of instances. The results appear to conform to contemporary theories on the development of representational use of symbols and specify the deleterious influence of long-term experiential deprivation on intellectual development.

The present study is part of a series investigating the role of language in the development of thinking in which deaf subjects have been observed in various situations designed to evaluate thinking in the absence of long-term linguistic stimulation (Furth, 1964; Furth, 1966). As this work has progressed, increasing use of Piaget's developmental system has been made; in particular, that part of his system which addresses itself to logical operations. This model has been found useful in describing subtle differences and similarities between deaf and hearing subjects with regard to the attainment of logical concepts.

Previously on this point, deaf adolescents, middle-class hearing subjects, and hearing adolescents from a rural, culturally restricted milieu were observed on tasks investigating separately the *discovery* and *use* of symbols standing for logical operations (Furth & Youniss, 1965). On the discovery problem the performance of middle-class subjects surpassed markedly that of deaf or hearing adolescents of a rural background in that only the middle-class subjects demonstrated through transfer that they had discovered the denotative properties of symbols for elementary logical connectives—'and', 'or', 'not'. In contrast, the task demanding the use of symbols, whose denotation was specified through instruction and training, failed to differentiate performance among these same groups. Successful symbol use equivalent to that shown uniquely by middle-class subjects on the discovery problem was now evident for all three subject-groups.

The form of this interaction between tasks and groups suggests the interesting possibility that cultural deprivation and linguistic deprivation produce similar effects in adolescents that differ from those produced in subjects afforded typical middle-

class experiences. Whether this finding has generality is one of the questions addressed by the present study. The major aim is to test for similarities and differences among groups of deaf, American Indian, and middle-class adolescents with regard to their use of symbols in specifying logical operations and their denotative properties. By reinvestigating the symbol *use* task, which previously yielded minimal differences, but at the same time employing younger subjects, extending training over several days and improving the method of assessing operativity were expected to increase the stability of performance and to make results more reliable.

The present study tested sets of concepts so that performance on any operation within a set could be related to success or failure on simpler component operations, thereby making an intra-subject analysis feasible. The two sets of concepts selected had one set dealing with the conjunctive operation 'and' and the other dealing with the conditional operation 'if. . .then'. These operations were chosen since they generate true and false instances of a complementary nature. That is to say, any simple conjunctive concept includes as true only one type of instance but many false types of instances, while conditional concepts yield only a single false-instance type but many types of true instances. This difference between one and many numbers of instances provides the opportunity to measure group differences as they might relate to breadth of concept or concept inclusiveness—factors considered relevant to the development of logical thinking (Inhelder & Piaget, 1964). Thus, at a minimum, the present sampling of concepts affords a possible systematic analysis of the ways in which deaf subjects and hearing but culturally deprived adolescents are alike in applying logical symbols according to logical rules.

METHOD

Subjects

Our sample of deaf children comprised twenty-four subjects enrolled in a state residential school for the deaf who fulfilled the following criteria: (*a*) hearing loss greater than 70 db in the better ear; (*b*) loss of hearing before CA 2; (*c*) no obvious accompanying behavioural or emotional problems; and (*d*) CA at least 114 months but not greater than 185 months (mean CA 145·5 months; S.D. 16·1 months). Also available for these deaf subjects were IQ scores derived from the performance scale of the WISC or WAIS. These tests were administered no more than eighteen months prior to the present experiment. The mean IQ for the group was 104·5 and the scores ranged from 86 to 131 with a median value of 105.

The culturally restricted subjects were twenty-four Indian children enrolled in a federally supported school on the Navaho reservation in Arizona. According to school records these children were reared in the typical Navaho rural setting on the reservation. Outside of school the boys mainly herded sheep while the girls either did this also or helped their mother care for younger children. Thus in the sense that these youngsters came from an environment which does not provide as much intellectual stimulation as would be found in a typical educated middle-class family, they are called 'culturally restricted' and thus, also, they resemble the rural sample in our prior study (Furth & Youniss, 1965). These children ranged in CA from 120–175 months (mean 138·9 months; S.D. 12·6 months).

The twenty-four control subjects were students enrolled in a public junior high school in a middle-class suburb of Washington, D.C. According to school administrators, most of the students came from families in which the father was employed in civil service positions within the federal government. These subjects ranged in CA from 143–170 months (mean CA 151·6 months; S.D. 6·8 months).

Procedure

Training. All subjects were seen in small groups in 30 min daily sessions for 5 days. During these sessions subjects were presented problems non-verbally on the blackboard and asked to volunteer to come to the front of the room to solve these problems. Each problem consisted of three parts: a symbolic expression (statement) on the left, a picture-instance drawn in outline on the right, and a truth value judgement in between; expressed → for true and ↛ for false. Problems were constructed by presenting two of three parts and responding consisted in subjects completing the missing portion.

Approximately three sessions were spent on the introduction of the method and the conjunctive operation 'and'. In the initial session the experimenter put on the board the following items and left them in sight of the class.

$H \rightarrow$ picture of a house	$R \rightarrow$ the colour red
$T \rightarrow$ picture of a tree	$B \rightarrow$ the colour blue
$S \rightarrow$ picture of a star	$Y \rightarrow$ the colour yellow
$A \rightarrow$ picture of an apple	

The presentation of a concept usually began with the writing of a statement which was true, then having children come to the board to construct other 'true' instances. This was followed by asking for 'false' instances, with the experimenter merely changing the truth value after a given statement. After this initial phase, problems were presented in which statement or truth value had to be filled in also. In addition to the filling-in procedure, two other techniques for training were used. The *elaboration* procedure required that subjects construct as many instances as they could for a given statement or as many statements as they were able for a given instance. The *contrast* technique presented one statement or one instance with both truth values so that true and false aspects of a concept could be viewed simultaneously.

From the items shown above a pool of instances was generated by combining a shape with a colour. These were changed arbitrarily from problem to problem so that each concept was presented in various ways. To illustrate, conjunction was shown as 'house and red' 'tree and blue', 'star and yellow', etc.

Concepts were introduced in the order: simple affirmation, simple negation, conjunction, exclusion, conjunctive absence, simple conditional, conditional with antecedent denied, and conditional with consequent denied. Affirmation required the presence of a single attribute—either a colour or a shape. Negation implied the absence of an attribute and was designated by a bar over the symbol standing for an attribute, e.g. $\bar{H}$ stood for the absence of a house. Conjunction always joined a shape with a colour and was designated by the symbol '$\bullet$' which appeared between two attribute symbols, e.g. $H \bullet R$ meant the presence of house and red. Exclusion combined the operations of conjunction and negation but was unique in postulating simultaneously the presence of one and the absence of another attribute. For instance, the statement $H \bullet \bar{R}$ referred to the presence of a house and the absence of red. The last concept in this set, conjunctive absence, also combined conjunction and negation but differed from exclusion in that it postulated the joint absence of two attributes. For instance, the statement $\bar{H} \bullet \bar{R}$ called for the absence of both house and red.

The final 2 days of training for most of the subjects consisted in presentation of the conditional operation, 'if . . . then'. Three concepts were used; the first was the simple conditional dealing with two affirmed attributes. For example, the statement $H \supset R$ required that if a house were present then it must be coloured red. The second concept was a conditional with the antecedent denied; $\bar{H} \supset R$. This statement said that if something was not a house it had to be coloured red. The final concept in this set was a conditional statement with the consequent denied. For example, $H \supset \bar{R}$ said that if a house were present it could be any colour except red. (Not all children in the present study were trained with the conditional concepts; some were trained instead with concepts not germane to this experiment.)

Testing. Either on the last (fifth) day of training or on the next school day subjects were administered a paper-and-pencil test to determine whether they could identify true and false instances of the concepts. According to currently used terminology this was a test of 'rule learning' or 'concept identification' for logical operations (cf. Haygood & Bourne, 1965).

Each item on the test presented a symbolic statement and a picture-instance; the subject had

to supply a truth value judgement. New shapes, not used during training, were introduced; these were M (man), C (car), and F (flag). All other symbols remained unchanged from those used during training.

Table 1 shows the composition of items on the test. Each concept was tested in four problems; two of these problems presented true instances and two problems false instances.

Table 1. *Description of concepts, symbolic expressions and instances used in testing*

(F, attribute flag; G, attribute green; f, picture of a flag; g, colour green; c, picture of a car; r, colour red.)

Concept	Expression	Truth value	Instance
Affirmation	F	True	f
		False	c
Negation	$\bar{F}$	True	c
		False	f
Conjunction	$F \bullet G$	True	fg
		False	fr
Exclusion	$\bar{F} \bullet G$	True	cg
		False	fg
Conjunctive absence	$\bar{F} \bullet \bar{G}$	True	cr
		False	fg
Conditional	$F \supset G$	True	cg
		False	fr
Conditional antecedent negated	$\bar{F} \supset G$	True	fg
		False	cr
Conditional consequent negated	$F \supset \bar{G}$	True	cg
		False	fg

Table 2. *Number of subjects making no error in two true and two false trials with five concepts*

		Concept and truth value														
		Affirmation			Negation			Conjunction			Exclusion			Conjunctive absence		
Group	n	T	F	Both	T	F	Both	T	F	Both	T	F	Both	T	F	Both
Control	24	22	20	19	20	20	18	17	14	12	19	13	12	20	18	16
Indian	24	21	16	15	5	15	4	13	16	11	14	11	8	9	12	5
Deaf	24	20	21	18	16	11	9	19	17	14	14	8	4	12	12	7

RESULTS

Single-attribute and conjunctive concepts

Table 2 reports performance for the concepts of affirmation through conjunctive absence as a function of subject-groups and truth value. For each concept, frequencies are first presented for the number of subjects who made no error in the two test trials with instances to be judged true and in the two trials with instances to be judged false. Ten χ^2 tests were made to evaluate differences among proportions performing errorlessly in the three subject-groups. Reliable differences resulted for negation with true instances ($\chi^2 = 20{\cdot}3$; D.F. 2; $P < 0{\cdot}01$) and with false instances ($\chi^2 = 7{\cdot}34$; D.F. 2; $P < 0{\cdot}05$) and for conjunctive absence with true instances

($\chi^2 = 11{\cdot}0$; D.F. 2; $P < 0{\cdot}01$). In each of these three cases the greatest proportion of successful subjects is found for the control subjects; least success in two cases was found among Indian children and in one case in the deaf sample.

In order to assess success with the overall application of concepts, scores were also computed for frequencies of subjects making no error in four trials with a concept. Since this criterion of success included performance with both true and false aspects of a concept, the scores in the third column under each concept indicate whether subjects can apply a logical rule completely, i.e. both to cases falling within and outside a rule. Of the five χ^2 tests made, reliable differences resulted for negation ($\chi^2 = 17{\cdot}1$; D.F. 2; $P < 0{\cdot}01$), exclusion ($\chi^2 = 6{\cdot}0$; D.F. 2; $P < 0{\cdot}05$) and conjunctive absence ($\chi^2 = 12{\cdot}0$; D.F. 2; $P < 0{\cdot}01$). In each case, control subjects succeeded in the greatest numbers compared to either Indian or deaf subjects. Thus congruence with previous results indicates that control subjects differ most notably from subjects in the other two groups on concepts which entail the operation of negation.

The final analyses of performance on this set of concepts considered the question of intra-subject consistency. If subjects were in fact correctly applying operations to identify instances, then performance on more complex concepts should be a function of success with a concept's simpler component operations. In order to assess this question three relations were considered: success on conjunction related to success on affirmation; success on exclusion; and success on conjunctive absence related to success on negation plus conjunction. First, a total of thirty-seven subjects among all subject-groups made no error in four trials with conjunction. Of these successful subjects thirty-one made no error also in four trials with affirmation. According to a two-tailed binomial test, the relation of success on these two concepts is reliable ($P < 0{\cdot}01$).

To measure success on exclusion and conjunctive absence performance on their components—negation and conjunction—was considered. An arbitrary criterion of success on the two components was set at a subject's making either one or no error in the eight trials testing these components. Of the twenty-four subjects who made no error in four trials with exclusion, twenty-three made one or no error in eight trials with its component operations. A two-tailed binomial test indicated a significant relation of success among these concepts ($P < 0{\cdot}01$). A similar result was obtained for success on conjunctive absence related to performance on negation plus conjunction. Of the twenty-eight subjects making no error on conjunctive absence, twenty-one subjects were observed to make one or no error on its components; the binomial test resulted in $P < 0{\cdot}01$.

Conditional concepts

Table 3 reports performance on the three conditional concepts for each subject-group; the first two columns present frequencies for subjects making no error in two true and in two false test items while the third column shows numbers of subjects who performed errorlessly in all four trials testing a concept. Only twelve of the twenty-four deaf subjects were trained and tested on these concepts (mean CA 154 months; range 139–165). Scores for both other subject-groups are given for the same twenty-four subjects whose performance appears above.

Proportion of success as a function of true and false instances was submitted to χ^2 tests assessing differences among groups of subjects. Two of the six tests yielded reliable results. For the true instance case of the simple conditional a $\chi^2 = 11{\cdot}5$, D.F. 2, $P < 0{\cdot}01$, was obtained; and for the true instance case of conditional with consequent denied, a $\chi^2 = 18{\cdot}2$, D.F. 2, $P < 0{\cdot}01$, was found. In both cases the greatest proportion of success was observed among control subjects, with smaller proportions observed among Indian and deaf children. Thus, with two of the three conditional concepts subject-groups were found to differ reliably in identifying true instances but in none of the three cases did a reliable difference result with false instances.

Table 3. *Number of subjects making no error in two true and two false trials with three conditional concepts*

		Concept and truth value								
		Conditional			Antecedent denied			Consequent denied		
Group	n	T	F	Both	T	F	Both	T	F	Both
Control	24	12	13	6	7	7	1	13	16	10
Indian	24	2	16	0	7	15	4	1	12	1
Deaf	12	2	3	0	4	7	3	1	4	1

Table 4. *Error scores for six trials as a function of operation and truth value*

		Conjunction		Conditional	
		True	False	True	False
Control	Mean	0·9	1·5	2·6	2·2
	Median	0·0	1·0	2·5	1·5
	Range	0–5	0–5	0–6	0–6
Indian	Mean	1·8	1·8	4·2	1·5
	Median	2·0	2·0	4·0	1·5
	Range	0–4	0–5	2–6	0–4
Deaf	Mean	1·5	1·8	4·0	2·6
	Median	1·0	1·5	4·0	2·5
	Range	0–4	0–5	2–6	0–6

No statistical tests were computed for success scores on all four trials of a concept. It is noted, however, that these data show a similar trend to frequency scores reported for two-trial performance. On the simple conditional and on the conditional with consequent denied concepts relatively more of the subjects in the control group performed errorlessly in four trials than in the other two groups. The opposite trend occurred with the conditional concept with antecedent denied on which the control subjects performed with relatively less success than the other groups.

The final analysis of performance was concerned with successful instance identification as a function of inclusiveness or breadth of concepts; inclusiveness is defined as the number of instance types included by an aspect of a concept. In the present system of logic there are four fundamental types of instances. Given two attributes simultaneously these are: (*a*) the presence of both attributes; (*b*) the absence of both

attributes; (*c*) and (*d*) the absence or presence of one attribute coupled with the presence or absence of the other attribute. All three conjunctive concepts—conjunction, exclusion, and conjunctive absence—yield true instances of a *single* type with the three other instance types being false. In complement, all three conditional concepts generate three types of true instances and only a single false instance type.

In order to perform analyses, a subject's errors were summed across concepts and truth values. As shown in Table 4 four scores were obtained for all subjects: (1) true instances of conjunction; (2) false instances of conjunction; (3) true instances of the conditional; and (4) false instances of the conditional. Each score was based on six trials. Simple sign tests were made to determine (*a*) whether a subject made more errors in twelve trials with the set of conjunctive concepts than in twelve trials with the set of conditional concepts, and (*b*) whether within a set of concepts a subject made more errors with true than with false instances. Results were: (*a*) consistently for control, Indian, and deaf subjects more errors were made with conditional than with conjunctive concepts ($P < 0{\cdot}05$ in each case); (*b*) for conjunctive concepts there was a trend for more errors to be made with false than with true instances but no reliable differences were obtained. For conditional concepts, Indian subjects and deaf subjects made significantly more errors with true than with false instances: $P < 0{\cdot}01$ and $P < 0{\cdot}03$, respectively. A non-reliable difference was noted for control subjects; $P > 0{\cdot}28$.

Discussion

Across all concepts, a nearly consistent pattern of differences was obtained among the three groups of subjects. Generally, when concepts or aspects of concepts denoted either a single instance or a single type of instance, the deaf, Indian, and control subjects did not differ on successful identification of instances. On the other hand, whenever symbolic expressions denoted a broader class of instances—such as with negated attributes or with true cases of conditional concepts—linguistically deprived deaf subjects and culturally restricted Indian subjects succeeded in significantly smaller proportions than normal control subjects. The answer to the initial question as to the similarity of performance between linguistically deficient and culturally deficient subjects is unambiguous. Performance patterns of deprived children are markedly alike and contrast with the pattern shown by middle-class children.

Most directly the results indicate differences in the ways control subjects and young adolescents of both of the other groups manipulate symbols standing for logical operations. As a first-level abstraction the data point to a distinction between using symbols literally to depict an instance and using symbols to represent a class of instances. Consider the fundamental difference between the operations of affirming and denying an attribute in the present system, as respectively depicting a single correct picture-instance *vs.* denoting any instance except the one denied. The relative success of all subjects with the operation of affirmation (see the concepts of affirmation and conjunction) in contrast to the relative failure of all but the normal controls with concepts involving negation suggests that when differences among subject-groups were observed, they occurred where symbolic statements referred to a broad class of instances rather than to single instances or a single instance type.

A related difficulty for deaf and for Indian subjects is apparent with two of the

three conditional concepts on which control subjects best succeeded in successfully identifying true instances. At the same time deaf and Indian children, but not the controls, showed greater difficulty in identifying true than false instances. The conditional operation, like negation, logically denotes a single false instance but positively it denotes a number of physically distinct but logically equivalent true instances. For example, the concept 'man implies yellow' denotes explicitly the one false case of *a man who is not yellow* while it includes as true the equivalent instances: *a yellow man*, *a yellow car*, and *a red flag*. Of course, only the first of these three true instances bears any literal similarity to the symbolic expression, while the latter two instances are denoted but not depicted by the expression.

The relatively broad inclusiveness of conditional concepts compared with the narrow and more literal inclusiveness of conjunctive concepts would account for the consistency with which conjunctive concepts are easier than conditional concepts across subject-groups. At the same time within conditional and conjunctive concepts the failure of aspects of these concepts to depict single instances led to poorer performance on the part of deaf and Indian subjects. Note that use of concepts that negate instead of affirm an attribute will not alone account either for group differences or relative difficulty among concepts. More precisely the number of instances which must be included under a concept appears to be the major factor best fitting present results.

As explanatory theory on these points, two recent views can be cited. Both Furth (1966) and Werner & Kaplan (1963) have referred to the developmental progression from depictive to referential behaviour correlated roughly with increasing CA. In brief, these authors agree that in the development of symbolic functioning a point is reached at which symbols no longer merely depict instances which bear a physical resemblance to one another, but they eventually come to refer to broader classes or events which can only be said to be included by the symbol.

Piaget also has addressed himself to the point at which symbols are no longer merely abstract depictors of reality but refer to logically equivalent but physically distinct instances. This developmental attainment is thought to occur at about CA 12, the age-range appropriate to the control subjects presently observed. In our previous investigations of children's attainment of various logical concepts (Furth & Youniss, 1965), CA 12 was also found to be the breaking-point at which symbols came to refer to classes or rules rather than to specific instances.

By implication, then, it is suggested that fewer deaf and Indian subjects than controls of the age-range presently sampled have yet reached the point in development at which this sophisticated use of symbols is attained. Note that in the control sample there is also clear room for improvement; success ranges from 80% on affirmation to 25% on the conditional. Perhaps with additional training there would have been less of a difference among groups. Nevertheless, it is not surprising that a greater similarity between control and deprived subjects was not found in the present study. Among other things, both deaf and Indian subjects have in common the fact that their educational instruction emphasizes the learning of English linguistic symbols with a strong stress on depictive labels. Whether this classroom emphasis or some more general environmental approach to dealing with symbolic content was operative cannot be decided.

It was expected that the methodological refinements and young age level employed in this study would help to differentiate long-term linguistic deprivation from cultural deprivation, a result which the prior study (Furth & Youniss, 1965) did not obtain. This expectation was not met. Nevertheless, the present data can be taken as evidence that mere acquisition of a natural language in no way guarantees that it will become an effective internal vehicle for intelligent symbolic functioning. If it did, then clear differences between Indian and deaf subjects would have been evident. Consequently the hypothesis that the linguistic deficiency of the deaf subjects has effects similar to experiential deficiency not accompanied by linguistic deficit (Furth, 1966; Furth & Youniss, 1965) is upheld by the present data.

This research was supported in part by grant no. 1484-S from the Vocational Rehabilitation Administration and by grant no. 02026–01 from the National Institute of Child Health and Human Development. Thanks are given to school administrators who kindly co-operated in our work, especially to the staffs of the deaf and the Indian residential schools.

References

Furth, H. G. (1964). Conceptual performance in deaf adults. *J. abnorm. soc. Psychol.* **69**, 676–81.

Furth, H. G. (1966). *Thinking without Language.* New York: Free Press.

Furth, H. G. & Youniss, J. (1965). The influence of language and experience on discovery and use of logical symbols. *Br. J. Psychol.* **56**, 381–90.

Haygood, R. C. & Bourne, L. E., Jr. (1965). Attribute- and rule-learning aspects of conceptual behaviour. *Psychol. Rev.* **72**, 175–95.

Inhelder, B. & Piaget, J. (1964). *The Early Growth of Logic in the Child.* New York: Harper & Row.

Werner, H. & Kaplan, B. (1963). *Symbol Formation.* New York: Wiley.

(*Manuscript received* 15 *May* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 445–448
Printed in Great Britain

THE CONTRADICTION OF EPIPHENOMENALISM: A RESPONSE TO E. A. LOCKE

By W. A. NORTON
*Department of Psychiatry, University of British Columbia**

In the past two years there have been contributions to this journal by Keehn (1964) and Burt (1964) on the topic of consciousness and behaviourism. In a further related paper, Locke (1966) has pointed out the logical difficulty inherent in any attempt to maintain the truth of epiphenomenalism, in so far as a behaviourist, for example, might be described as believing this doctrine and, at the same time, concerned with establishing its truth. The argument appears to the present writer to be somewhat contrived. It may be that there are other grounds on which behaviourism is open to criticism as a would-be exclusive movement in psychology than simply the apparent contradiction of epiphenomenalism as examined from the perspective of those who hold other views about consciousness.

Behaviourism as a method

As Ryle (1949) has observed, there is some analogy between the intrinsically private nature of consciousness and what he regards as the historically antecedent religious tenet concerning the privacy of conscience. Because of this peculiarity of consciousness, particular philosophical problems have always attended the attempted verbal examination of 'mind', or subjective experience, and its supposed special relationship to what people say and do. Such problems become further compounded, however, in moving from one's own experience to speculate upon that of others. The possible variations from person to person in respect both of the quality of their experience as well as their functional capabilities become of concern. Indeed, some of the earliest researches in the nascent scientific psychology—the studies of Galton, for instance—dealt with both sorts of individual differences. Later, however, what were seen as insuperable difficulties in the use of traditional introspective methods for the study of thought processes, together with the desire to follow the objective methods of the physical sciences, led to the emergence of the behaviourist movement. This started off with some promise of giving an objective account, at least of human learning and development. At that time, the questions of the actual content of any human learning, and of the possibly significant dependence of learning upon the social roles of teaching and controlling were irrelevant as far as psychology was concerned. What was new, incidentally, was not so much the method, as its would-be application to psychology as a whole, and the emphasis on the method—which emphasis led in turn to a tendency towards procedural preoccupation. Be this as it may, against this background of the emergence of behaviourism as a method in psychology, the ascription to behaviourists in general of epiphenomenalist beliefs (a conceptual contradiction, as Locke insists) seems illicit within their adopted terms of

* Now at Department of Reform Institutions, 434 University Avenue, Toronto 2, Ontario

reference. If not an argument 'ad hominen', it might be described as projective or animistic, specifying the latter term, in this instance, as attributing to others, i.e. to behaviourists, a consciousness, including knowledge, ideas, opinions, beliefs and concern with truth, similar to the critic's own. In fact, the philosophical doctrine of epiphenomenalism with its various implications is somewhat to one side of behaviourism in psychology.

Locke (1966) quotes Burt (1964) as pointing out that epiphenomenalism makes nonsense of the work of educators, moral reformers and others. However, such activities appear to proceed at least as successfully as elsewhere in countries whose prevailing ideologies are usually held to entail epiphenomenalism. It is one thing to question the validity of these exhortations. Roughly two-thirds of subjects may respond to a variety of teaching techniques, for example, or the same proportion learn the material spontaneously. It is a different question, however, whether or not such activities look nonsensical from the point of view of a non-behaviourist.

Causal influence and determinism

The aspect of the doctrine that Locke elaborates and which is important for his argument, namely, that ideas and opinions or statements thereof are without mutual influence, perhaps does not present a formidable obstacle to the epiphenomenalist. The process of education or any such suasive activity may surely be described in terms of physical stimulation which, in accordance with its nature as a field process in a physiological system, is more or less internally consistent or congruous. Indeed, Locke's argument regarding the nexus between the alleged ideas and the public statements of epiphenomenalists (he cites behaviourists) seems to beg the traditional so-called body/mind problem. As he perceives, the invocation of determinism is an answer in itself within the epiphenomenalist's terms of reference; and it is to introduce quite other considerations and terms of reference to raise such issues as that of man being 'free to judge and to judge their judgements' (this sort of thing leads to an infinite regress); and making 'a valid claim to truth'. It is worth observing, incidentally, that another deterministic position, that of psychoanalysis, far from decrying consciousness, places particular value on this aspect of human activity.

How behaviourists behave

Locke also raises the question of how behaviourists behave in respect of the beliefs which he imputes to them; the behaviour here, of course, referring to their formulation of theories about behaviour and their statements regarding them. If extended beyond this juxtaposition the argument might go on to show that behaviourists adopt a whole repertory of behaviour patterns as time and occasion demand. These include not only professing behaviourism, but also real life behaviour as a teacher, parent and so on, in accordance with the socially accepted and changing rules and principles governing these roles, and which seem to demand more than S-R theories for their adequate description. Of these sorts of behaviour and the rules, we do not usually postulate beliefs on the part of the person concerned. Behaviourists and others, indeed, appear to share in an implicit behavioural 'psychology' whereby they adapt with varying success to the biosocial environment, to some extent regardless of what

they may do, say or even think about behaviourism. Any cognitive strain which their versatility may engender is a matter of private concern.

Truth and the well

What underlies the different points of view of epiphenomenalists and opposers of this view is perhaps a variety of personality factors, in a broad sense, including both cognitive factors as well as the usual objective—tough-minded—extravert, as opposed to the subjective—tender-minded—introvert dichotomy. Could such an hypothesis be investigated without the methodology imposing its own answer, however, or without reducing the problem to the experimental dimensions appropriate for those who talk of taking a problem into the laboratory? It is of peculiar moment that method selects the content of research, and of theory. If introspection and preoccupation with consciousness lead, on the one hand, to untrammelled thought, with or without images; so behaviouristic studies seem mainly to deal with directing and controlling behaviour within a demarcated and artificial environment. Despite the often abstract generosity of their titles, these studies do not usually permit much wider generalization. This may be true of other than behaviourist studies, and indeed criticism of much psychological game activity has recently been made by Dunnette (1966). But whether or not there are, or whatever the nature of, such differences between psychologists who adhere to behaviourism and those who do not, neither logical nor empirical investigation is likely to produce agreement about the truth of this matter, or even to clarify the issue while one party refuses to be bound by a sort of case law of philosophical precedent which it disowns. The purpose, after all, was to change things which hitherto had not lacked philosophical speculation.

Perhaps what should and does concern behaviourists, who after all are not notable exponents of epiphenomenalism, is the rate at which things are, in fact, changing. Some, for example, Broadbent (1961), seem to regard behaviourism's half-century as having shown adequate progress. Where, however, some application of learning theories has been begun say in programmed instruction or behaviour therapy, it is not yet clearly established if these methods are more efficacious and efficient, or what the relationship is between their effect and the theories on which they are based.

Within this admittedly arbitrary context of applied utility, however, and in respect of epiphenomenalism as a doctrine rather than of behaviourism as a method, one might question also whether or not the promotion of deterministic theories along with the adoption of more numerous or modified means of influencing behaviour is altogether an advantageous choice or strategy, and so not merely a matter of historical inevitability (Berlin, 1954). Such a question is only meaningful in a society and climate of opinion that regards itself as 'open' in Popper's sense. It seems expedient, therefore, as a number of writers have urged, including Burt, Broadbent, and Richardson (1965), that psychology should not confine itself exclusively to the behaviourist approach.

References

Berlin, I. (1954). *Historical Inevitability*. London: Oxford University Press.
Broadbent, D. E. (1961). *Behaviour*. London: Eyre & Spottiswoode.
Burt, C. (1964). Consciousness and behaviourism. A reply. *Br. J. Psychol.* **55**, 93–6.
Dunnette, M. D. (1966). Fads, fashion and folderol in psychology. *Am. Psychol.* **21**, 343–52.

KEEHN, J. D. (1964). Consciousness and behaviourism. *Br. J. Psychol.* **55**, 89–91.
LOCKE, E. A. (1966). The contradiction of epiphenomenalism. *Br. J. Psychol.* **57**, 203–4.
RICHARDSON, A. (1965). The place of subjective experience in contemporary psychology. *Br. J. Psychol.* **56**, 223–32.
RYLE, G. (1949). *The Concept of Mind.* London: Hutchinson.

(*Manuscript received* 5 *October* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, 449–450
Printed in Great Britain

THE NATURE OF EPIPHENOMENALISM: A REPLY TO E. A. LOCKE

By MAURICE J. A. GLICKMAN*
Department of Sociology, Hebrew University, Jerusalem

In a recent issue of this *Journal* Edwin A. Locke (1966) seeks to refute epiphenomenalism on the grounds (*a*) that the doctrine destroys the integrity of the epiphenomenalist's own statements, and (*b*) that it is based on a doctrine of psychological determinism which is also self-contradictory. Unfortunately, his argument is based on a misunderstanding of the logical and philosophical nature of epiphenomenalism.

Epiphenomenalism seems to be understood in different ways by different thinkers, but for the sake of convenience Locke's own definition may be quoted: 'Epiphenomenalism is the doctrine that all mental contents, for example, ideas, values, opinions, emotions and the like, are an incidental by-product of physiological or physical processes and that they have no causal efficacy either in influencing behaviour or in influencing any other ideas the individual has.'

Locke first urges that if the doctrine were true then the behaviourist's own behaviour in stating it would have no logical or empirical connexion with his own thoughts on the matter; a conclusion which Locke describes as absurd. But of course the doctrine only excludes a causal relation between an idea and any of the individual's activities which follow it. It does not exclude the possibility that both events may be determined by the same or similar (physical) forces, and indeed it is this very implication which Locke attacks in the second part of his argument. As he correctly observes, epiphenomenalism clearly implies psychological determinism, the doctrine that all human thoughts and actions are caused. If this latter principle, he claims, were applied to the epiphenomenalist's own position then the epiphenomenalist would have to admit that his belief in psychological determinism was itself determined by previous conditions. If so, 'how can he claim that what he is asserting constitutes knowledge or truth?'

It is here that Locke's misunderstanding of his adversary's behaviour lies. A behaviourist, if he does assert the doctrine that all psychological events are predetermined (he normally *implies* it) rarely claims that it is true. There is a subtle but important difference between asserting a belief and claiming that it is true. An assertion is an expression of a (dogmatic) assumption, which itself may or may not be true but on the basis of which a great deal of knowledge or truth may be acquired. Nobody rejects mathematics because of such notions as the square-root of minus one; nor, because Newton's Laws have never been tested, nobody ever having observed a body in a state of complete rest or uniform motion, do we refuse to accept any results of empirical research in physics which utilizes classical mechanics.

What is here suggested is that epiphenomenalism is an attitude, and not, strictly, an hypothesis. Like other basic assumptions in science its value depends not on its own foundations but on the results of inquiry into nature which are based on it. In

* Now at Department of Anthropology, Brandeis University, Waltham, Mass. 02154.

fact there is no reason why, accepting the premise of epiphenomenalism, psychologists should not be able to investigate the physiological dimensions of any human thought or action, as long as opportunities are practicable.

Cause in physics is a concept which is used to express relations between physical events separated by time. Physicists cannot explain the coming into existence of the whole of the physical universe, since to do so would be to invoke another physical event which is itself a part of the physical universe as a whole. The concept of cause has also failed to fully explain the motion of individual electrons within the atom. But neither of those facts deters physicists from exploiting causation to its limits. Similarly, psychologists of a behaviourist inclination ought not to be dismayed at the fact that they cannot explain the physical basis of all thought processes because of the methodological difficulty that the understanding of every thought process entails a further thought process.

It remains to point out that Locke's argument against epiphenomenalism is of the same logical order as epiphenomenalism itself. His also is a philosophical attitude which can make no special claim to truth. There are several ways of showing this. Let one suffice. It is generally accepted that if one side in a dispute assumes from the start, in however disguised a form, the falsity of its opponent's position, then its argument may be very persuasive, but it will fail to be logically convincing. This is precisely the error, which is in the nature of any argument against epiphenomenalism, which Locke makes. From the beginning he implies that the argument that he will present will be a completely objective evaluation which we should not regard as expressions of some physical or physiological reality. But what he is trying to *prove* is that objective evaluations are possible. In the end his proof turns out to be based not on the examination of evidence, but on a premise which it covertly contained all the time.

Ultimately the test of the fundamental principles of science is of a pragmatic rather than a logical or empirical order. Time only will tell whether epiphenomenalism will possess the significance for psychology such as Newton's Laws possessed for mechanics or whether it will turn out to be just another passing dogma.

REFERENCE

LOCKE, E. W. (1966). The contradiction of epiphenomenalism. *Br. J. Psychol.* **57**, 203–4.

(*Manuscript received* 16 *February* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 451–453
Printed in Great Britain

EPIPHENOMENALISM AND THE CONTRADICTIONS OF ITS CRITICS

By H. J. EYSENCK
Institute of Psychiatry, University of London

Locke (1966) takes up the cudgels from Burt (1964) to castigate the treatment which consciousness receives from behaviourists. Unfortunately his argument is marred by logical errors which reduce its impact to a *non sequitur*.

He argues that: 'if the behaviourist *could not help* believing and saying what he says, that is, if everything he thought and did were determined by forces outside his control, then how can he claim that what he is asserting constitutes knowledge or truth? If everything he does is determined by forces outside his control, the same must apply to his espousal of the theories of epiphenomenalism and psychological determinism. No proponent of psychological determinism can claim to have reached any conclusion on the basis of objective evidence which he is free to judge and evaluate; all he can assert is that he was forced to believe in and to assert his theories by his past conditioning, his childhood environment or the like. If men are not free to judge and to judge their judgements and to reach conclusions based on their assessment of the facts of reality, then no conclusion reached by anyone could make a valid claim to truth, including the theory of psychological determinism'.

In this argument, note first that the term 'truth' is question begging. Scientists do not use the term in Locke's sense at all; we are under no delusion that our theories 'make a valid claim to truth', but adopt a much more pragmatic test, e.g. that they give rise to testable deductions which are in fact verified (in some cases) by experimental evidence. Philosophers used to make assumptions about the 'eternal truth' of their more speculative pronouncements, but this practice has fallen into desuetude; it is difficult to see why Locke should wish to revive it. Pragmatic interpretations of 'truth' certainly do not require consciousness for their acceptance; a computer could be programmed to assess successful and unsuccessful consequences of actions, such as those encountered in laboratories, just as they have been programmed to deduce consequences from theoretical statements.

Note next the importance Locke attributes to psychological functions, such as believing; 'how then is the behaviourist to explain the fact that *he does believe* the theory to be true and at the same time *does behave* as if he believes it to be true, i.e. that he asserts it at all and that there is some empirical connexion between his beliefs and his public statements?' I certainly do not believe my theories to be true; I should say that they have a certain heuristic value, give rise to testable predictions, and will no doubt in the near future be supplanted by better theories. There is no epistemological difficulty to this; my attitude to the theory is governed by the reinforcement received (to put it crudely) in the matter of verification. The more successful predictions are made from the theory, the more firmly will I become attached to it; Skinner's well-known experiment on superstition in pigeons exposes

the fundamental basis of this process, as well as containing a warning against taking such theories too seriously, even when apparently successful.

How then about the main point of Locke's statement, to wit, '. . . that the doctrine (of epiphenomenalism) makes nonsense of the argument for epiphenomenalism and of anything its proponents might say'. If anyone did indeed claim that this doctrine was true in some absolute sense, one would have to agree that he was talking as much nonsense as anyone propounding the opposite claim. But considered as a heuristic theory, leading to certain testable consequences and verified by certain observations, it does not seem to be more nonsensical than alternative theories, and possibly less so; too few experimental observations are perhaps available as yet for it to be possible to hold any very strong views. But on the grounds of alleged 'contradictions' the doctrine must be absolved; unless we start by assuming the truth of theories opposed to epiphenomenalism, it may be regarded as perfectly self-consistent.

It should hardly be necessary to answer Burt's point that 'such a doctrine makes nonsense of the work of the educator, the psychotherapist, the moral reformer, the legislator and the historian. . .' (Burt, 1964). The educator wants to produce a given result, say that of making his pupils speak colloquial French with ease and in good Parisian accents; the psychologist designs a modern language laboratory to test different methods for achieving this goal, on the assumption that behaviour is predictable, and that certain antecedent conditions will have certain consequences. How does that 'make nonsense' of the work of the educator? The therapist wants to rid his patient of his facial tic, or his phobia, or his alcoholism; he uses the methods of negative practice, of reciprocal inhibition or of aversion therapy, based on behaviouristic principles as verified in the laboratory. His success does not seem, to me at least, to 'make nonsense' of the work of the psychotherapist; it seems, for the first time, to make it possible for the psychotherapist to do in fact what his patients have for so many years begged him to do, and which he was unable to do. In any case, a theory is not untrue because it 'makes nonsense' of cherished beliefs; these beliefs are simply irrelevant.

Throughout Locke's paper runs the apparent belief that epiphenomenalism and determinism are the same, or imply each other, and that the alternative to determinism is some form of 'free will'—whatever that might be. Both beliefs are wrong. One can be an epiphenomenalist without believing in determinism, or a determinist without believing in epiphenomenalism; there is no essential logical connexion. And the alternative to determinism is not free will. Heisenberg's principle, which has sometimes been adduced by opponents of determinism as disproof of that doctrine, does not carry any implications of free will, as a look at the mathematical statement rather than the usual verbal paraphrases will make abundantly clear; at most it opposes chance or chaos to complete determinism. It is doubtful if this alternative will please exponents of free will any more than does determinism. Free will is not a meaningful alternative because it is not a doctrine which can be stated in a logically consistent and experimentally testable form; it is a poetic expression of human aspiration, not a scientifically useful concept (Eysenck, 1961).

The confusion implicit in the notion of free will is expressed most clearly in the quotation from Locke (1966) given in the first paragraph of this reply, where he claims that only if men are free to judge and reach conclusions based 'on their assessment of

the facts of reality' will their conclusions be in accord with truth. But surely their judgements and conclusions are constrained precisely by these 'facts of reality', and they are also constrained by attitudes, preferences, and personality factors which are partly inborn and partly acquired through environmental pressure. The former factors are useful in assessing the heuristic value of theories, whereas the latter probably bias any such assessment; in either case it is difficult to see what meaning could possibly attach to any claims that people are 'free' to judge and to assess the facts of reality.

In summary, epiphenomenalism is not a very good theory, but it is the best we have. Determinism is a useful heuristic doctrine, acceptable until there is some good evidence which forces us to give it up. Neither is likely to be 'true' in any absolute sense, and neither involves us in logical contradictions of the kind suggested by Locke. Fortunately all this is not very important; the relation between philosophical preconceptions and fruitful scientific work is weak at best, and probably non-existent.

References

Burt, C. (1964). Consciousness and behaviourism. A reply. *Br. J. Psychol.* **55**, 93–96.
Eysenck, H. J. (1961). The concept of mind: a behaviouristic approach. *J. psychol. Res.* **5**, 1–11.
Locke, E. A. (1966). The contradiction of epiphenomenalism. *Br. J. Psychol.* **57**, 203–4.

(*Manuscript received* 15 *November* 1966)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 455–458
Printed in Great Britain

A NOTE ON TRENDS IN ANSWERS TO CATTELL PERSONALITY QUESTIONNAIRES BY SCOTTISH SUBJECTS

By JOHN McQUAID
Child Guidance Centre, Saltcoats, Scotland

Answers to personality questionnaires by eight groups totalling 1733 subjects from south and south-west Scotland show marked trends towards 'anxiety' and less definite tendencies towards 'introversion' by comparison with American norms for the tests. The result for introversion confirms that of a previous study of British subjects by Cattell & Warburton (1961), but the result for anxiety is idiosyncratic and requires confirmation by other studies on large samples. Cattell's 'children's', 'high school' and '16 PF' personality questionnaires were used.

Table 1, below, summarizes the results of testing children, adolescents and a small number of adults with Cattell personality questionnaires over a period of 5 years. Despite the variety of the eight groups of over 1700 subjects, some well-defined trends appear, centred mostly on high anxiety, and more ambiguously on introversion, relative to American norms. A comparison is also made with the results of the study of British and American students by Cattell & Warburton (1961).

Cattell (1957, p. 418) referred to greater introversion among British subjects compared with Americans, and to higher than average scores on 'surgency' (F+) among French subjects. Cattell & Warburton (1961) again found British students to be significantly more introvert and less anxious than their American counterparts. Americans were also less 'radical' (Q_1). With regard to the alleged dourness of Scots, it happens that Cattell (1957, pp. 269–70, 892) attempted a definition of this complex of traits; he described it in test terms as shrewdness, restraint and pessimism, together with some sign of 'schizothyme temperament' (A−).

Subjects. The 1700 subjects fall into eight widely divergent groups. For this reason, and also because of uncertainty as to three normal requisites for averaging (complete identity of factor from age-group to age-group; comparable variance; and normality of score distribution) these groups have been kept separate throughout. The large number of delinquents and the small number of adults reflect the type of assessment normal in the writer's experience of child-guidance work. All of the testing was done by the writer with the help of staff. The non-delinquent cases are from the south-west of Scotland, and the older delinquents from approved schools in this area and, to a lesser extent, in south and east Scotland. There is no certainty that the 320 controls (group 5) for these delinquents (group 6) exactly match the latter, though the consistency of score patterns and the comparisons with the other groups strongly suggest that, except for low intelligence in the delinquents, the divergence is no greater than would be expected between well-matched groups. It is also not unlikely that delinquents are pre-selected for certain socio-cultural factors which still evade accurate analysis and definition. The total population tested numbers over 2000, of whom about one-third had been referred for child-guidance treatment as abnormal in any way. The small group of primary-school teachers (group 8) volunteered to take the '16 PF' test anonymously.

The following is a brief account of the eight groups of subjects:

Group 1 (n = 320). Primary-school pupils: boys and girls aged 8–14 yr (tested in schools or at child-guidance centres, but not referred for treatment on account of abnormal behaviour).

Group 2 (n = 90). Delinquents: boys aged 8–14 yr (tested in schools or at child-guidance centre, after being placed on probation).

Table 1. *Mean factor scores (and* S.D.*) for answers to Cattell personality questionnaires by eight groups of Scottish subjects, in cases where the mean factor scores differ reliably from the American norms*

Group	*n* (Sex)	Subjects (age in years)	Test†	First-order factors‡								Second-order factors	
				C−	D+	H−	O+	Q_2+	Q_3−	Q_4+	F−	Anxiety	Introversion
1	320 (M and F)	Primary pupils (8–14)	CPQ (A)	.	2·06** (1·04)	2·19** (1·03)	.	.	.	.	.	.	.
2	90 (M)	Delinquents (8–14)	CPQ (A)	.	2·52** (1·09)	.	.	.	.	.	.	.	.
3	30 (M and F)	Clinical (8–14)	CPQ (A)	.	.	.	.	.	.	.	.	.	.
4	300 (M and F)	Primary pupils (8–13)	CPQ (A 1, A 2)	5·05** (1·39)	5·81** (1·43)	.	.	.	5·09* (1·82)	.	.	5·8	.
5	320 (M and F)	Secondary pupils (13–18)	HSPQ (A)	5·04** (2·45)	.	.	5·99** (2·13)	6·21** (2·11)	5·10** (1·82)	.	.	5·8	5·7
6	600 (M)	Approved school (11–18)	HSPQ (A)	.	5·90** (1·58)	.	6·10** (1·86)	.	5·08** (2·30)	6·42** (1·99)	.	5·9	5·7
7	33 (M)	Seminary students (13–17)	HSPQ (A)	4·06** (2·45)	7·06** (1·85)	.	6·61** (1·40)	.	.	6·52** (1·95)	.	.	.
8	40 (F)	Primary-school teachers	16 PF (A/B)	.	.	4·30** (2·35)	.	7·08* (3·47)	.	.	4·30** (2·12)	6·0	6·4

* $P < 0·05$. ** $P < 0·01$.

† CPQ, Children's Personality Questionnaire (forms A, A1 or A2); HSPQ, High School Personality Questionnaire (form A); 16 PF, Sixteen Personality Factor Questionnaire (forms A and B).

‡ The factor scores for groups 1, 2 and 3 are 'raw' factor scores (corresponding American norms: mean = 0, S.D. = 1·0); in all other cases the factor scores are sten scores (corresponding American norms: mean = 5·5, S.D. = 2·0).

Group 3 (n = 30). Clinical cases: 21 boys and 9 girls aged 8–14 yr (referred to the child-guidance centres as maladjusted). This group did not differ significantly on the factors included in Table 1.

Group 4 (n = 300). Primary pupils: 184 boys and 116 girls aged 8–13 yr (tested in schools).

Group 5 (n = 320). Senior and junior secondary pupils: 175 boys and 145 girls aged 13–18 yr (tested in schools).

Group 6 (n = 600). Boys in eight approved schools in Scotland (tested in classes).

Group 7 (n = 33). Students aged 13–17 yr, in a Scottish junior seminary. The entire age-group was tested in class.

Group 8 (n = 40). Primary-school teachers employed in Ayrshire schools.

Results

No score has been entered for a group unless the score of every subject in the group for the factor in question was on the same side of the mean established as the norm for American subjects, i.e. 5·5 or over for plus scores, and 5·5 or under for minus scores. In this way a definite meaning is ascribed to the word 'trend'. It also happens that, apart from intelligence, for which delinquents in Scotland invariably make a low score, no significant score is excluded by this condition. Thus significant, qualifying scores appear in approximately one-sixth of the possible cases, i.e. in about 20 cases out of the possible total of eight groups by 14–16 factors. Raw scores have been considered a more useful basis of comparison for the first three groups, since transformation to sten scores entails too great a relative sacrifice of precision. In the other cases the American sten means (fixed at 5·5) and S.D. (fixed at 2·0) are the basis of comparison.

The significant factors, all of which are constituents of either anxiety or introversion in test terms, are listed below. They are given in order of importance.

Factor D+ (excitement high) is significant for about 1300 subjects. It covers excitable, demanding, impatient, attention-getting behaviour. Excitability is a Cattellian constituent of anxiety.

Factor O+ (high guilt feeling) is significantly high for over 1200 cases. A feature of the anxiety syndrome is that it describes worried, anxious, depressed, remorseful behaviour.

Factor Q_3- (self-sentiment low) appears significantly low in almost 1000 subjects. The definition covers uncontrolled emotionality, and rejection of cultural demands. It is an important contributor to anxiety in test terms.

Factor C− (ego strength low) features in about a third of the sample. Low scores are associated with emotional instability, annoyability, low frustration tolerance and emotional immaturity. It is an important factor in anxiety.

Factor Q_4+ (ergic tension high) is central to anxiety, and characterizes irrational worry, tension, turmoil and irritability.

Factor Q_2- is a constituent of introversion in Cattell's terms, and describes resolute, decisive behaviour and group-resistant opinions.

Factor H− (threat reactivity). Autonomic over-responsiveness reveals itself in a shy, withdrawn, aloof, basically schizothyme temperament, and hence has a high hereditary determination. It is an element in the second-order factor of introversion.

Factor F− (desurgency) summarizes glum, sober, introspective, value-orientated behaviour, and is an important contributor to introversion.

The only remaining constituent of introversion not included in the table of results is factor A− (schizothyme rather than cyclothyme temperament); the weighted mean for 1300 subjects was 5·1. This completes the picture of introversion and anxiety.

The second-order scores for anxiety and introversion have been derived by com-

bining separate constituent factor scores by the methods described in the handbooks for the various tests (Cattell *et al.*, 1959–63). The high scores for the anxiety syndrome do not agree with the values obtained by Cattell & Warburton for British (English) students; but the introversion trend is concordant with their results (see Cattell & Warburton, 1961, Table 2, p. 10). In addition, the primary-school teachers show very high 'radicalism' (Q_1+), not included in the table, which replicates the findings of Cattell & Warburton for British students on this factor.

REFERENCES

CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement*. New York: World Book.

CATTELL, R. B. *et al.* (1959–63). *Handbooks to the Children's, High School and Adults'* (16 *PF*) *Questionnaires*. Champaign, Ill.: I.P.A.T.

CATTELL, R. B. & WARBURTON, F. W. (1961). Cross-cultural comparisons of patterns of extraversion and anxiety. *Br. J. Psychol.* **52**, 3–15.

(*Manuscript received* 21 *April* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 459–462
Printed in Great Britain

A NOTE ON FIELD DEPENDENCY AND MOVEMENT AFTER-EFFECTS

By E. BAKER and N. F. DIXON
University College London

While it has been maintained (Goldstein, 1958) that the spiral after-effect is a special case of the 'waterfall' illusion, the very much greater strength of the latter, when evoked by a rotating field of randomly distributed black and white squares (Dixon & Meisels, 1966), suggests that the underlying processes may differ in the two cases.

Phenomenally both the spiral and the random field after-effect have two components—a gross movement of the entire test field coupled with a viscid creeping of its microstructure. Inasmuch as apparent movement of the test field as a whole depends upon integration of the movements of the elements of which it is composed, it might be argued that the two phenomenal components reflect, respectively, contributions from both central and peripheral levels of the visual system. If this is so then it might further be supposed that the retinal contribution would be proportionately very much greater in the case of the random field than in that of the spiral figure, since retinal transmission rate as a function of interface displacements within the retinal image is much greater in the former case.

The investigation reported here examined this possibility by testing the hypothesis that central determinants of perceptual experience will have a relatively greater influence upon the spiral after-effect than upon a movement after-effect induced by a random display of black and white squares.

The central determinant chosen was the trait of 'field dependence' (Witkin, Dyk, Faterson, Goodenough & Karp, 1962). The prediction was that subjects who are relatively more field dependent would show proportionately greater after-effects for a spiral as opposed to a random stimulus field when compared with subjects who are less field dependent. The rationale was that, while all subjects would show greater after-effects for the random than for the spiral field, field-dependent subjects would be more susceptible to centrally determined components of such after-effects.

Method

Subjects. The experiment was carried out on two groups of subjects, field-dependent and field-independent as follows:

	Field dependent	Field independent
Male	6	6
Female	3	6

These classifications were based on results of an application of the Hidden Figures Test, Form V (as described by Jackson, Messick & Myer, 1964, and obtainable from the Educational Testing Service, Princeton, New Jersey) to a larger group of twenty-five males and twenty-two females. The scores of those selected for this investigation were above and below the top and bottom quartiles of the scores obtained.

Apparatus. The stimulus fields consisted of two 8 in. diameter disks, one displaying a three throw spiral and the other a quasi-random distribution of black and white squares. At a viewing distance of 3 ft. the disks subtended an angle at the eye of 12° 30′. Both disks were driven by a 'Velodyne' motor set to give a rotation speed of 2 revs/sec for the spiral and 7 revs/min for the random field. The test field, a grey screen, was positioned at right angles to the plane of the disks and at the same distance from the subject. A second screen was used to conceal whichever of the two disks was not being used in any one trial.

Procedure. Having been shown the two disks and seated in front of them, each subject was given the following instructions: 'Watch the centre of the disk as it revolves and continue watching until I say "now". When I say "now" turn your head and look at that grey screen. You will find that the part of the screen you are focusing on will appear to rotate if you have been looking at the random field, or appear to expand or contract if you have been watching the spiral one. I want to determine the length of time that this after-effect lasts, so during each trial tell me as soon as you no longer see any movement on the screen.'

Under the above procedure, six readings, three for clockwise and three for anticlockwise rotation, were taken for each subject on each disk, the order of presentation being counterbalanced, with half the subjects having the spiral disk first and the other half the random disk first. The initial viewing period for both disks was 30 sec.

The durations of the after-effects were measured by stop-watch.

RESULTS

The results for all subjects are given in Table 1, including the means for various subgroups. The following observations may be made:

1. Nineteen of the twenty-one subjects showed longer movement after-effects for the random than for the spiral field. This finding is significant (binomial test, $P < 0{\cdot}001$).

2. The main hypothesis that field-dependent subjects would show a smaller difference between the random and spiral after-effect than would field-independent subjects is not borne out.

3. Sixteen out of twenty-one subjects showed a longer-lasting expanding after-effect after clockwise rotation than they did a contracting after-effect after anticlockwise rotation of the spiral. This difference, in agreement with results from previous researches (see Holland, 1966), is significant ($P < 0{\cdot}01$, two-tail). Perhaps more surprising is the fact that thirteen subjects showed more prolonged after-effects following anticlockwise as opposed to clockwise rotation of the random field. This result is again significant (Wilcoxon test, $P < 0{\cdot}02$, two-tail).

4. There was the further finding that presentation order affected the illusion for field-dependent subjects. For these subjects only, the random field after-effect was significantly greater when it preceded rather than when it followed the presentation of the other field ($P < 0{\cdot}05$).

DISCUSSION

As they stand the results seem to lend no support to the initial hypothesis that the spiral after-effect would be more influenced than the random field after-effect by a central determinant of perceptual experience. However, an examination of the data indicates that a monolithic prediction about the relationship between field dependence and the spiral after-effect was inappropriate. The female subjects showed the predicted relation between field dependence and duration; but among the males there was a marked tendency for short spiral durations to be associated with field depen-

Table 1. *Duration (sec) of movement after-effects for 'random squares' and spiral stimulus disks: means of six responses for each subject*

		Random squares		Spiral		Means		Differences
	Subject	Clockwise	Anti-clockwise	Clockwise	Anti-clockwise	Random squares	Spiral	(Random – spiral)
Field-dependent males	1	25·8	31·4	20·9	11·0	28·6	16·0	12·6
	2	17·8	11·5	8·8	6·6	14·6	7·7	6·9
	3	9·6	11·1	0·0	3·4	10·4	1·7	8·7
	4	2·1	1·5	2·2	2·0	1·8	2·1	−0·3
	5	17·8	20·6	9·4	5·5	19·2	7·4	11·8
	6	15·9	20·9	0·0	0·0	18·4	0·0	18·4
Means		—	—	—	—	15·5	5·8	9·7
S.D.		—	—	—	—	8·2	5·4	5·7
Field-dependent females	1	19·3	13·8	11·6	10·5	16·5	11·0	5·5
	2	10·3	10·0	9·8	3·3	10·2	6·6	3·6
	3	6·2	7·3	12·3	9·0	6·8	10·6	−3·8
Means		—	—	—	—	11·2	9·4	1·8
S.D.		—	—	—	—	3·9	2·0	4·0
Field-independent males	1	19·7	17·9	18·4	11·5	18·8	15·0	3·8
	2	16·8	16·3	12·8	9·2	16·6	11·0	5·6
	3	22·2	34·3	22·0	16·9	28·3	19·4	8·9
	4	18·2	17·8	8·9	0·0	18·0	4·5	13·5
	5	13·5	19·4	16·9	12·1	16·4	14·5	1·9
	6	11·6	15·3	8·1	9·1	13·5	8·6	4·9
Means		—	—	—	—	18·6	12·2	6·4
S.D.		—	—	—	—	4·6	4·7	3·8
Field-independent females	1	10·8	11·8	5·5	0·0	11·3	2·8	8·5
	2	19·0	22·4	18·2	14·4	20·7	16·3	4·4
	3	11·5	11·2	5·2	5·7	11·4	5·4	6·0
	4	15·5	22·7	7·2	7·0	19·1	7·1	12·0
	5	12·3	12·8	11·8	4·7	12·5	8·2	4·3
	6	11·2	12·9	7·6	10·7	12·0	9·1	2·9
Means		—	—	—	—	14·5	8·2	6·4
S.D.		—	—	—	—	3·9	4·1	3·0
Means (males)		—	—	—	—	17·0	9·0	8·0
Means (females)		—	—	—	—	13·4	8·6	4·8
Means (all subjects)		14·6	16·3	10·4	7·3	15·5	8·8	6·7

dence and long durations with field independence. A possible explanation is that an additional correlate of field dependence was operating in the situation. The subjects' task was a decision-making one, and Kogan & Wallach (1964) found that sex acts as a moderating variable on the association between field dependence and risk-taking behaviour. They reported that field-dependent males tend to be conservative, while field-independent males take more risks. For women the reverse of these relations holds true. If the risk-conservatism variable was of importance in this experiment, and if our hypothesis that the spiral figure would be more susceptible to central effects than the random figure is true, then one would expect the difference scores (random – spiral) to be smaller for field-dependent women than for field-independent women, but to be larger for field-dependent men than for field-independent men.

Reference to Table 1 shows this to be the case. It therefore appears that a more effective test of our hypothesis would have to exclude risk-taking behaviour from the experimental procedure.

REFERENCES

DIXON, N. F. & MEISELS, L. (1966). The effect of information content upon the perception and after effects of a rotating field. *Q. J. exp. Psychol.* **18**, 310–18.

GOLDSTEIN, A. G. (1958). After effects of the 'waterfall' and 'spiral' illusions. *Am. J. Psychol.* **71**, 608–9.

HOLLAND, H. C. (1965). The spiral after effect. *Int. Ser. Monogr. exp. Psychol.* vol. 2.

JACKSON, D. N., MESSICK, S. & MYER, C. T. (1964). Evaluation of group and individual forms of embedded-figure measures of field dependence. *Educ. psychol. Measmt.* **24**, 177–92.

KOGAN, N. & WALLACH, M. A. (1964). *Risk Taking—a Study in Cognition and Personality.* New York: Holt, Rinehart & Winston.

WITKIN, H. A., DYK, R. B., FATERSON, H. E., GOODENOUGH, D. R. & KARP, S. A. (1962). *Psychological Differentiation.* New York: Wiley.

(*Manuscript received* 16 *January* 1967)

Br. J. Psychol. (1967), **58**, 3 *and* 4, *pp.* 463–465
Printed in Great Britain

SIMPLE APPARATUS FOR MULTIDIMENSIONAL VIBROTACTILE SIGNALLING

By D. D. DIESPECKER
University of Newcastle, N.S.W., Australia

Investigation of the skin as a communication channel (for transmitting speech) was attempted in the 1920s and 1930s, particularly in Germany and the U.S.A., but it was abandoned because, 'no-one paused to ask the skin what language it could compass' (Geldard, 1957).

The human integument is sensitive to different forms of energy: mechanical thermal, electrical, and chemical. Pressure as a form of communicable energy has been investigated systematically and this has led to detailed research into mechanical vibratory stimulation. This mode is one of two which are currently receiving most attention in this field, the other being electrical stimulation.

Cutaneous communication has special relevance for those handicapped in one or more sense modalities and also for those living or working in unusual environments, e.g. for man in space and for man beneath the sea. Although there exists a variety of techniques and apparatus for vibrotactile and electrocutaneous communication, there is a scarcity of published descriptions of such apparatus. The apparatus described here had the advantages of being simple and of being economical to construct. It may be built quickly and has the added advantage of being robust. It has performed without malfunction in more than 120 hr of experimentation (Diespecker, 1967).

The equipment: design requirements

The requirements of the system were:

(*a*) To transmit forty coded elements of information (the letters of the alphabet, ten digits, and four frequently used words) from a console to five electromechanical transducers.

(*b*) To transmit eight signals to each of the five transducers or vibrators in such a way that any one vibrator would produce 'hard' 'medium', and 'soft'; and single, double, or treble vibrations when energized and attached to the skin of a subject.

(*c*) To allow future conversion of the console so that a flip-flop circuit and relays could be added and the single, double, and treble signals replaced by signals having three durations of time, namely, 0·1, 0·3 and 0·5 sec.

The control unit

The control circuit is shown in Fig. 1. The $\frac{1}{2}$ A transformer provides a choice of four voltages (10 V, 15 V, and 20 V proved satisfactory as the three degrees of intensity used to operate vibrators in a programme of experiments which allowed subjects to learn coded signals with high accuracy). The switches are of the spring-loaded doorbell type and are screwed on to a hardboard or aluminium panel. Power may be taken from a domestic mains outlet. Since vibrotactile experiments, particularly learning experiments, usually require signals to be transmitted one at a time to a particular vibrator, the control unit contains five outlets. Each switch represents one element of information and the experimenter should avoid transmitting more than one signal at a time in order to avoid overloading the coils (and confusing his subject).

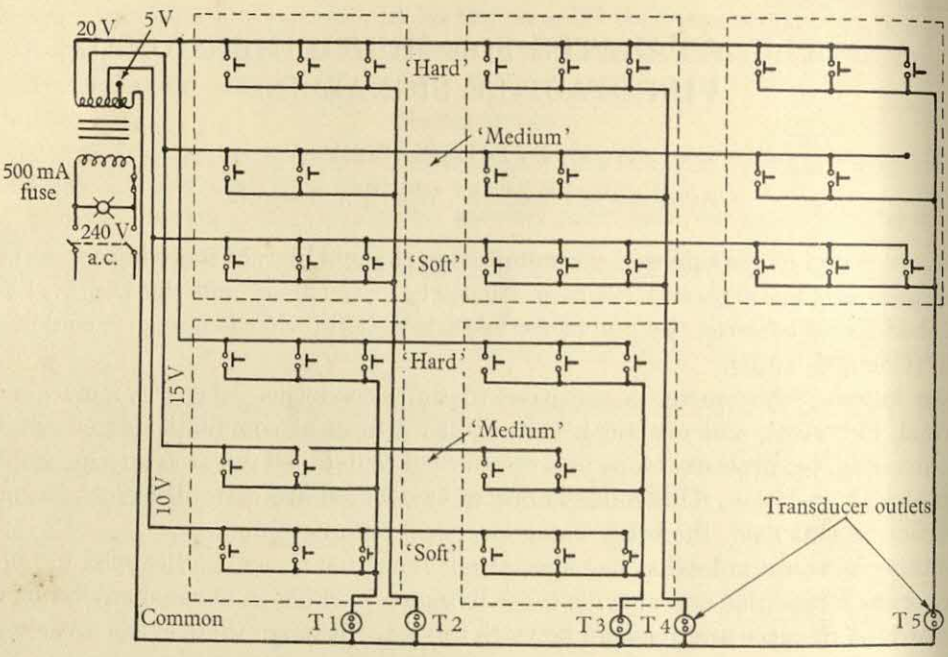


Fig. 1. Circuit diagram for control unit.

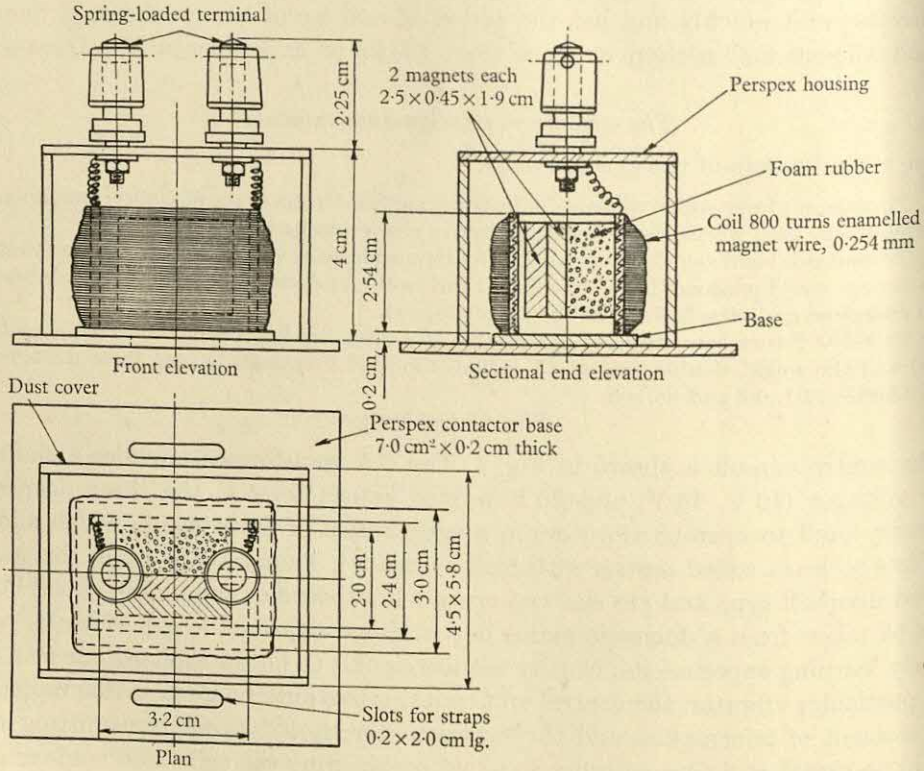


Fig. 2. Electromechanical vibration transducer.

Electromechanical vibration transducer

A transducer (vibrator) is shown in Fig. 2. Five of these units are plugged into the control unit and each is attached, by means of elastic strapping and snap fasteners, to a particular body area, e.g. arms and legs. The original design was by Sherrick, who has described the method of construction (Sherrick, 1965). This unit differs in detail from Sherrick's in terms of wire diameter, terminals, and contactor base area. The vibrator has a resistance of approximately 50 Ω and an impedance, at 50 c/s, of 50 Ω.

Operation

The experimenter may operate the switches at a convenient distance from his subject so as not to distract him by movement and to prevent his subject from seeing the console. The system works satisfactorily when the vibrotactile code described by Howell (1956) is used. For convenience, the particular code in use may be marked on each switch. In the Howell system the encoding of letters and digits on the keyboard is based on the frequency of occurrence of letters used in English (as is the Morse code). Vowels, for example, are represented, one to each vibrator, by single 'hard' (20 V) pulses and letters having a lower frequency of occurrence such as Z or J are encoded as two 'hard' pulses. Since frequently used letters can be transmitted faster than infrequently used ones (either by single rather than multi pulses, or by single short pulses of various durations in more sophisticated apparatus) a variety of useful communication systems can be designed.

When the experimenter closes a switch on the console one of the vibrators is energized and the entire unit vibrates until the switch is released. Vibrations from any one unit are thus transmitted to any object to which it is attached. The vibrations can neither be seen nor heard by a subject in an experiment.

References

Diespceker, D. D. (1967). Vibrotactile learning. *Psychonom. Sci.* **9**, 107–8.

Geldard, F. A. (1957). Adventures in tactile literacy. *Am. Psychol.* **12**, 115–24.

Howell, W. C. (1956). Training on a vibratory communication system. (Unpublished Master's thesis, University of Virginia Library.)

Sherrick, C. E. (1965). Simple electromechanical vibration transducer. *Rev. sci. Instr.* **36**, 1893–4.

(*Manuscript received* 5 *April* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 467–476
Printed in Great Britain

'BRAIN AND CONSCIOUS EXPERIENCE': A CRITICAL NOTICE FROM THE U.S.S.R. OF THE SYMPOSIUM EDITED BY J. C. ECCLES (1966)*

By A. R. LURIA
University of Moscow

In recent years there has been an upsurge of interest in the problem of the relation of consciousness to the brain and to underlying brain mechanisms. At least four large international symposia have been devoted to this problem since 1953 (Adrian, Bremer & Jasper, 1954; Jasper *et al.*, 1958; Wolstenholme & O'Connor, 1960; Moruzzi, Fessard & Jasper, 1960).† This paper is devoted to the fifth and latest in the series (Eccles, 1966).

There are many reasons for such an interest in this problem. It is partly due to recent developments in brain morphology connected with the discovery of the non-specific reticular formation of the brain and to the rapid development of physiological research at the neuronal level. It is also due to recent successes in neurosurgery which have made it possible to conduct experiments in which the human cerebral cortex is stimulated while the subject is at the same time being questioned. There is again the recent development in the study of memory and its analysis at the molecular level. Finally, there are the recent achievements in psychopharmacology which have made it possible to follow the changes in consciousness brought about by different centrally acting drugs. But perhaps the basic reason for such a lively interest in the relation of the brain to conscious activity is the effort in natural science to do away with the 'isolation' of the mental and the attempt to try to understand the material basis of conscious activity.

The symposium on *Brain and Conscious Experience* occupies a special place in the efforts to discuss this problem. It is distinguished by two characteristics. First, perhaps no other symposium dedicated to the problem of the relation of brain to mind has brought together such a large number of outstanding participants, including such scholars as Adrian, Bremer, Eccles, Granit, Jasper, MacKay, Moruzzi, Mountcastle, Penfield, Phillips, Sperry, Teuber and Thorpe. These scholars represent the fields of neurophysiology, morphology, psychology, ethology, information theory and cybernetics. This character of the symposium facilitated a broad and many-sided discussion of the problems.

A second characteristic of the symposium was the fact that it was called by the Vatican Academy of Sciences and took place in one of the Vatican palaces. Pope Paul VI addressed the participants, pointing out the significance which the Roman Catholic Church attaches to scientific research on the brain and its relation to consciousness.

The symposium was preceded by a correspondence between its organizer, Sir John

* *Brain and Conscious Experience*. Edited by J. C. Eccles. Berlin, Heidelberg and New York: Springer Verlag. 1966. Pp. xxi + 591. $16.80.

† A symposium on the problems of consciousness was also held in Moscow in the spring of 1966.

Eccles, and the President of the Vatican Academy of Sciences. In this correspondence the President of the Academy asked that a purely scientific character be attached to the symposium and that discussion of philosophical questions be excluded. In reply Eccles declared that in his opinion all science has a philosophical basis, that 'there is a philosophy of science which is in fact basic to all scientific investigations' and that 'it is not possible to exclude relations with philosophy' (p. viii).* The outcome was that philosophers were not invited to participate, but that philosophical questions nevertheless continued in fact to be very important.

Basic philosophical positions

Almost every participant in the symposium began with a theoretical definition of consciousness and the characteristics of his attitude to the philosophical problem of its relation to the brain. Hence it is of serious interest to analyse the basic philosophical positions of these representatives of modern science.

A careful study of the material from the symposium clearly shows the position which has developed in the natural sciences in modern capitalistic countries. The real achievements of modern science and those scientific findings which objectively are essentially materialistic differ sharply from the philosophical positions of the researchers themselves. Although their factual work is armed with the most up-to-date of techniques and is conducted on the highest and most modern level, their initial philosophical positions remain behind the factual work and display the typically theoretical helplessness of idealistic philosophy. This is why, in accordance with Wundt's 'principle of the heterogony of ends', the factual achievements of these scholars go far beyond the boundaries of their formulated aims and enrich science in spite of the fact that the philosophical problems they set before themselves lead them nowhere.

The philosophical positions of a significant number of the participants in the symposium (Adrian, Eccles, Granit, Penfield and others) are not distinguished by anything novel or original. They define consciousness as a primary, directly given reality and consider the objective world as being a secondary reality produced from consciousness. Basing themselves on this dualistic principle (in which they repeat the position of the classical physiologist, Charles Sherrington), they set themselves the basic task of precisely defining the relation of consciousness to brain, of finding those brain mechanisms in which 'consciousness enters the brain' or in which brain units begin 'to generate consciousness'.

Several of the leading participants in the symposium (for example, Eccles), basing themselves on the principle that 'the primary reality is our consciousness—everything else is a derivative and has a second-order reality' (p. 327)—took the position of 'methodological solipsism' (p. 315). Others (for example, Granit) 'belong to the people who were brought up in the Machian philosophy' (p. 255). Many cited Sherrington, Eddington and such modern representatives of idealistic philosophy as Beloff (1962), Kneale (1962) and Wiegner,† in whose statements they see the basis for the indisputable admission of dualistic philosophy and refer to the idea that 'we

* The page references throughout are to *Brain and Conscious Experience*, the work under review.
† E. Wiegner, 'Two kinds of reality', unpublished lecture, 1964.

must retain the Platonic notion of mental events which are distinct from anything in the physical world' (p. 472).

It is quite natural that such positions led many of the participants to consider as fundamental the question posed by Sherrington, namely in which formations of the brain does the 'physical give rise to the mental' and in what way does active volition which can result in our hands shading our eyes from the sun produce the influence of the mental on the physical? (See, for example, pp. xv, 315, 446–7.)

One can only be thankful that the factual work of these participants goes well beyond the limits of these questions. Being guided by the logic of objective scientific research, they have enriched science with many facts of outstanding scientific value. There were also participants, MacKay, Moruzzi, Teuber and others, who were distinguished from the others by theoretical positions which have nothing in common with the above conceptions of subjective idealism.

Two scientific positions

In spite of the fact that the definitions of consciousness given by the participants were varied, not one of them understood consciousness as the reflexion of objective reality, as 'conscious being' or as complex activity which has a semantic and systemic structure. In what way then do these investigators approach the analysis of the material substrate or brain mechanisms of consciousness? In their approach to this basic question the participants were clearly divided into two groups.

Some of the participants consider that one should look for the material basis of consciousness *inside the brain* and that careful search will lead to the discovery of formations in the neuronal structure of the brain which give rise to consciousness. Among these is Penfield who, on the basis of observations made at the operating table when he has stimulated the human cortex electrically, has come to the conclusion that, in addition to the sensory and motor zones of the cortex, there exists a region which when stimulated calls forth experiences and thoughts. This area, limited to the temporal zone, he considers to be the 'interpretative' cortex which in its earlier stages of development was neutral (see 'The uncommitted cortex', pp. 217–36). There are also those neurophysiologists like Eccles who do not consider it possible to be limited to a crude macroscopic level in the search for the mechanisms of consciousness. They believe that these mechanisms can only be found on the neuronal and synaptic level. Following this idea, Eccles returned to the notion he expressed earlier (cf. Eccles, 1952): that the elementary forms of the phenomenon of consciousness ought to be sought by the methods of quantum physics, by seeking out the smallest synaptic units at a level where the Heisenberg–Eddington 'uncertainty principle' can be replaced by a principle of harmonic organization; he again repeated the idea that the basis is the synaptic unit, the mass of which is expressed by the figure 10^{-16} g (p. 467). Such explorations of the nervous basis of a non-qualitative consciousness are far removed from the study of the mechanisms of consciousness understood as the complex meaningful reflexion of reality.

Among the researchers who were trying to solve the problem of the material basis of consciousness in the above way were neurophysiologists working on the neuronal and molecular level. However, the representatives of the schools of information theory and cybernetics had a completely different position on the question of the

mechanisms of consciousness. MacKay claimed with justice that it was unlikely that the problem of consciousness could be solved by the study of isolated neurons. In this he received support from researchers who claim that neurons in any part of the nervous system have an identical structure (see Eccles, p. 50) and that, consequently, it is not possible to obtain facts which will lead to the mechanisms of consciousness from the study of separate neurons.

Consciousness according to MacKay is a form of organization of a whole system concerned with the processing of information, and its mechanisms should be sought in the properties of organizations of this system. Regarding the transfer of information as something 'psychophysically neutral', and repeating in this a well-known contention of idealistic philosophy, he sees in the study of the structure of the system which transmits information 'the bridge which connects the data of physiology with conscious experience'. On this basis he considers it possible to explain not only the phenomenon of consciousness but also 'free will' without disturbing the principles of the structure of physical systems (pp. 427–9, 433–5). The position of the biophysicist, Gomes, was similar; from an analogous position he attempted to approach a solution to the problem of how one should understand the possibility of transforming the uncertainty which reigns at the quantum level into the certainty characteristic of conscious activity (pp. 449–59).

Although the approach of the first group of researchers leads far away from the analysis of those large and complex structures permitting complex forms of reflexion of reality, it still retains the possibility of concrete physiological research. The approach used by the representatives of the second group of scholars, while retaining whole complex systems as a subject of analysis, radically departs from any concrete research and replaces the analysis of real physiological mechanisms by the 'black box' principle of description. There are obvious dangers connected with this approach in which scientific research is replaced by logical systems.

Thus two scientific positions were presented at the symposium; they are characteristic of modern science and potentially they contain the rudiments of a serious crisis in the future.

New tendencies in research on brain systems

Discussions of fundamental theoretical positions did not, however, constitute the major part of the symposium. The larger part was devoted to the analysis of material reflecting essential progress in science. This material was very varied. There were papers reflecting the current state of morphological and morphophysiological studies of the cortex and the brain stem (Colonnier, Eccles and Anderson); papers on the analysis of the physiological mechanisms lying at the basis of sensory and perceptual processes (Mountcastle, Granit, Creutzfeldt, Libet and Teuber); and papers dealing directly with the problem of brain mechanisms of consciousness and wakefulness (Penfield, Adrian, Jasper, Eccles and Moruzzi). There were also papers containing new facts about brain mechanisms involved in the unity of the personality and reports of experiments in which different regions of the brain have been separated (Bremer and Sperry), on the cerebral control of movement (Phillips), and papers concerned with attempts to approach cerebral organization in the light of contemporary cybernetics (MacKay and Gomes), as well as papers in the field of ethology (Thorpe).

It is not possible to consider each of these topics separately or to reflect the richness of all the factual information contained in these reports. Only some basic questions are considered, therefore, and some of the trends in brain research which are reflected in the symposium.

Non-specific and specific systems of the brain

The central problem of the conference on 'Brain mechanisms and consciousness' in 1953 was the question of the role played by the non-specific system of the brain stem in providing an active and waking state for the cerebral hemispheres (Adrian *et al.*, 1954). The discovery of the reticular formation of the brain stem and the optic thalamus made by Magoun, Moruzzi and Jasper, and the neurological analysis of brain-stem mechanisms by Penfield attracted widespread attention. One might say that the 1950s were to a significant extent years of research in neurophysiology on the non-specific systems of the brain, and that research on modally specific cortical systems was set aside for the time being.

Now the situation is essentially different. The analysis of the neuronal structures of the cortex and their vertical connexions, and neurophysiological investigations of the functions of separate neurons, connected with the development of microelectrode techniques and the introduction of computers, have brought about fundamental changes. The work of Hubel and Wiesel, Jung, Evarts and many others has shown the variety of ways in which neurons differ from one another, and what highly specific functions neurons have for perceptual processes. After Jung had shown that there are neurons which respond to stimuli of various modalities, as well as neurons which regulate only one modality, a group of investigators has shown that there are neurons in the cortex and in the corresponding nuclei of the thalamus which respond exclusively to such specific stimuli as smooth or broken, vertical or inclined lines, movement from the centre of the field of vision to the periphery and from the periphery to the centre, and so on. Thus it has been established that information reaching the sensory cortex is broken down into many thousands of component elements in order to be synthesized dynamically in the more complex areas of the cortex. It has also been established that, together with the cortical neurons (or subcortical nuclei) which react to different aspects of a stimulus, there are neurons which react only to a *change* in the perceptual field (the 'novelty neurons' of Jasper, or 'attention neurons' of Hubel, and of Sokolov). Furthermore, it has been established that the populations of these neurons have a sufficiently clear topographical distribution within the cerebral hemispheres, forming well-defined zones in which neurons of a particular type predominate in vertically organized 'colonies' (Mountcastle and others). These zones form distinct structures within the cortical and subcortical apparatus of the brain. Other new evidence shows that formations can be distinguished in the cortex in which 'novelty neurons' predominate (e.g. hippocampus) and which evidently are connected with such complex functions as 'memory'. Finally there is the careful research which has correlated the character and density of cell populations in different regions of the cortex with their functional characteristics and their 'graded accessibility' to different stimuli, for instance (Phillips and others).

All these facts have stimulated a growth of interest during the past decade in

'specific' formations of the brain and in a return to a new level of what Mountcastle calls the 'geographic principle' in brain study (p. 86). Granit (p. 122–3), Jasper (p. 272) and several other participants in the symposium supported this view. It led the participants categorically to reject Lashley's idea of the equivalence of different regions of the cortex, an idea which has been shown to be the result of imperfect experimental technique on Lashley's part (see Thorpe, p. 552; Teuber, p. 553; and others). They emphasized that the non-specific reticular system itself acts in close connexion with the specific systems of the brain stimulating them and receiving stimulation from them. (P. K. Anokhin has also worked on this problem in the U.S.S.R., although his work was not reported at the symposium.) They emphasized that it is more realistic always to have in mind the interrelations of the non-specific (brain-system) and specific (cortical) formations which together are responsible for the real forms of regulation of behaviour (Jasper, p. 272). This rebirth of the 'geographic principle' of brain analysis, on a new plane, is one of the basic features of the symposium.

Selectivity in the work of brain systems

The positions formulated above reflect current approaches to the morphology or 'statics' of the brain; but there was also great interest in new approaches to the *dynamics* of the brain system.

Recent decades have seen a rejection of the idea of simple transmission of information along nerve tracts leading to the central nervous apparatus and the storage of new and rich facts. Instead it seems that the reception of information by the nervous system is an *active* process and that at practically all levels there is a *selection* of necessary, 'useful' information, accompanied by inhibition and blocking of unnecessary, non-essential stimulation. This increased interest in ideas which were formulated in their day by I. P. Pavlov in his theory of 'analysers', and of analysis and synthesis as the basic function of the cortex, is very apparent in the symposium, although Pavlov's ideas were not themselves mentioned.

A group of participants put the view that the essential function of the nervous apparatus is not so much the transmission of excitation due to external stimulation as the *limitation* of the excitation transmitted, the selection of the essential and the inhibition of the superfluous (Mountcastle, p. 89; Moruzzi, p. 556; Granit, p. 117; and others). This process is connected with the independently operating processes of continuous excitation in the cortex and with synaptic inhibition (Eccles, p. 24). All these facts have turned up in the observations of ethologists, who have shown that animals react only to a relatively limited number of signals which have been selected in the history of the species and retained genetically in the form of 'innate releasing mechanisms' (Thorpe, p. 475), or which have been selected by experience in the lifetime of an individual. These facts have turned up as well in a series of neurophysiological investigations which show that even separate neurons can be activated only in the presence of conditions corresponding to a specific 'aim' (Granit, pp. 129–32; Moruzzi, p. 354) and that, to quote Moruzzi, 'in life we receive only those impressions which can be utilized' (p. 557).

Facts concerning the state of continuous excitation of neurons in regions of the brain (Jasper, pp. 260–1) do not say anything about 'the spontaneous work of the cortex' as some of the theoreticians feared. Rather, these facts show that in the

nervous apparatus there can develop continuous dominating foci which participate in the selection of needed information. The emergence of such foci brings us closer to the analysis of memory mechanisms. Jasper (p. 562) states that with the development of an epileptic focus there arises in the symmetrical region of the other hemisphere a 'mirror focus' which can be retained for months after the expiration of the basic epileptic zone. It is of interest that, as Moruzzi suggests (pp. 349–53), 'these slow continuous processes appear first of all in the most complex neurons of the cortex which participate directly in the digestion of information and the elaboration of new forms of experience'. Their establishment requires the long sleep which is necessary for complex neurons but which is completely unnecessary for more simple motor neurons. In this way sleep spreads throughout the cortex in a highly selective fashion (p. 439). The analysis of the mechanisms of this selective activity naturally drew the attention of researchers to the results of synaptic studies, and Eccles' report (pp. 24–54, 321 ff.) reflects accomplishments in this area of investigation.

The recognition of the active selective character of the central nervous apparatus is one of the most significant features of modern neurophysiology and it receives much attention in the symposium.

Functional systems of the brain

Classical physiology as a rule was characterized by investigations of sensory and motor functions in isolation and by the analysis of the morphology of receptor and motor zones. The methods used were stimulation of the cortex and recording of potentials in the cortical regions corresponding to the sensory or motor periphery. Such investigations have created serious doubts. One of the basic tendencies of current psychophysiology is to develop the study of whole functional (sensori-motor) systems. This tendency is clearly in evidence in the symposium.

Mountcastle and Teuber on the one hand, and Phillips on the other, show in their papers that both the sensory and the motor apparatus of the cortex work in much more intricate dynamic complexes than had been assumed, and that they operate according to much more complex laws than had been previously described. As Phillips has shown (pp. 391–2) the stimulation of one isolated point of the sensory cortex never leads to a whole sensation nor stimulation of an isolated point of the motor cortex to a whole action. The effects of such stimulation always have an artificial character and it gives rise only to parts of the corresponding sensory or motor process. Actually, motor processes are affected by much broader systems of cortical zones, the significant part of which goes far beyond the limits of the 'motor' cortex. Thus there is new support for the idea that afferentation widely distributed across the cortex participates in the organization of movements. (This idea has been advanced by L. A. Orbeli, P. K. Anokhin, N. A. Bernstein and others in the U.S.S.R.)

Investigation by Phillips has uncovered interesting new facts concerning the intimate organization of the motor zones of the cortex. It has been known for a long time that stimulation of separate points in the cortex leads to results which change greatly depending on the intensity of the stimulus and the preceding state of the cortex.

Phillips (pp. 401 ff.) has succeeded in showing the presence in the motor cortex of 'optimal points' of excitation corresponding to places of densest population of motor

cells and in formulating a principle of the selective accessibility of separate muscle groups. According to this principle, the muscles of those zones which possess the greatest mobility and directability (fingers of the hand, lips, tongue) have not only the greatest representation in the motor zone of the cortex (which has been convincingly shown by Penfield), but also have the densest population of motor cells in the cortex, and thus the greatest 'selective accessibility' of directable impulses from the periphery (pp. 339–402). This fact requires a re-evaluation of the brain mechanisms of motor processes and careful analysis of those afferent mechanisms of the cortex which make them possible.

Phillips and other participants introduced the especially interesting idea that among those types of structures which can provide the directability of movements there are also structures created by internal 'efferent patterns' which permit great regulation of a single efferent impulse, even when it is isolated from sensory systems (pp. 408–9). Such facts greatly change the customary understanding of the role of afferent systems in the regulation of movements and they deserve intensive study.

Teuber is among those who have thought it necessary to investigate whole functional systems underlying perception and movement. His paper is concerned with perceptual disturbances in brain lesions. Facts about perceptual disturbances in local brain lesions, and disturbances of constancy in particular, led him to conclude that human perception cannot be considered to be the result of the work of a sharply delineated zone of the cortex. The whole system of brain zones participates in every perception—visual perceptions for instance—and each zone makes its contribution in coding visual information. Some zones are concerned with the reception of visual information; others are concerned with the spatial organization of that information; and a third group is concerned with constancy. Of great interest is the fact that, according to Teuber, in perceptual constancy which is retained independently of 'groping' movements of the perceiving organs (in particular the eyes), an active part is played by those returning impulses which come from the movements to the perceiving apparatus. When analysing functional systems, therefore, one ought not to go only from sensory to motor, as is usually done, but also from motor to sensory (pp. 198 ff.). This principle, which was supported by MacKay (pp. 429–31), provides a new approach to the formation of active reflexion which was outlined by I. M. Sechenov in his day and which has been developed in the U.S.S.R. by N. A. Bernstein.

Bremer (pp. 283–95) and Sperry (pp. 298–308) introduced new material on the study of functional systems of the brain which make possible the unity of personality. These papers present new facts concerning the various mechanisms which make possible the joint working of both hemispheres and describe in detail those amazing forms of disturbances of the unity of psychological activity which were seen in Sperry's widely known experiments. Fibres of the corpus callosum in monkeys (and now also in man) were completely severed, isolating one hemisphere from the other and transforming them (both the dominant one concerned with speech and the subdominant one which is not concerned with speech) into two completely independently working organs. These observations are of great interest to psychology and merit intensive study.

The change from isolated studies of the function of separate regions of the brain

to the careful study of functional systems, of kinds which have been carried out in the U.S.S.R. by the physiologist P. K. Anokhin and the psychologist L. S. Vygotsky, represents an important tendency in the development of current psychophysiology which was clearly reflected in this symposium.

Conclusion

This symposium is one of the most significant that has been devoted to the problem of the relation of the brain to conscious experience. Physiological science, having rid itself of its one-sidedness, has begun to arrive at a synthesized understanding of nervous processes underlying psychological activity and has brought out facts concerning the wholeness of the relations between the non-specific and specific systems of the brain, and the selective operation of the brain in reflecting reality; it is now approaching an objective analysis of the workings of functional systems. All these scientific achievements will receive new directions of development when the problem of the relation of the brain to consciousness is posed on a correct theoretical basis. Physiology now has the task of uncovering the mechanism of consciousness. Consciousness is intelligible, not as an insubstantial subjective state, but as achieving the most complex form of the reflexion of reality by means of its equipment of semantic and schematic structure.

The author gratefully acknowledges the help of Mr Norman Segalowitz, who translated the typescript into English.

Selected Bibliography

Adrian, E. D., Bremer, F. & Jasper, H. H. (eds.) (1954). *Brain Mechanisms and Consciousness*. Oxford: Blackwell.

Anokhin, P. K. (1935). *Problems of Centre and Periphery in the Physiology of the Higher Nervous Activity*. Moscow: Gorki. (In Russian.)

Anokhin, P. K. (1955). New data on the afferent apparatus of conditioned reflexes. *Vop. Psychol.* no. 6. (In Russian.)

Banshchinov, V. M. (ed.) (1966). *Problems of Consciousness*. Moscow: Acad. Sci. (In Russian.)

Beloff, J. (1962). *The Existence of Mind*. London: MacGibbon and Kee.

Bernstein, N. A. (1935). Problems of the interrelations of co-ordination and localization. *Arch. biol. Sci.* **38**, no. 7. (In Russian.)

Bernstein, N. A. (1947). *On the Construction of Action*. Moscow: State Medical Publishing House. (In Russian.)

Bernstein, N. A. (1967). *The Co-ordination and Regulation of Movement*. London: Pergamon.

Eccles, J. C. (1952). *The Neurophysiological Basis of Mind*. Oxford: Clarendon.

Eccles, J. C. (ed.) (1966). *Brain and Conscious Experience*. Berlin, Heidelberg and New York: Springer-Verlag.

Jasper, H. H., Proctor, L. D., Knighton, R. S., Noshay, W. C. & Costello, R. T. (1958). *Reticular Formation of the Brain*. Boston, Mass.: Little.

Kneale, W. (1962). *On Having Mind*. Cambridge University Press.

Moruzzi, G., Fessard, A. E. & Jasper, H. H. (eds.) (1960). *Brain Mechanisms—Progress in Brain Research*. Amsterdam: Elsevier.

Orbeli, L. A. (1949). *Problems of the Higher Nervous Activity*. Moscow–Leningrad: Acad. Sci. (In Russian.)

Pavlov, I. P. (1949). *Complete Works*. Moscow–Leningrad: Acad. Sci. (In Russian.)

Sechenov, I. M. (1952). *Physiology of the Nervous Centres*. Moscow: Acad. Med. Sci. (In Russian.)

Vygotski, L. S. (1956). *Selected Psychological Works*. Moscow: Acad. Pedagog. Sci. (In Russian.)

VYGOTSKI, L. S. (1960). *Problems of the Development of Higher Psychological Functions*. Moscow: Acad. Pedagog. Sci. (In Russian.)

WOLSTENHOLME, G. E. W. & O'CONNOR, N. (eds.) (1960). *The Nature of Sleep: CIBA Foundation Symposium*. London: Churchill.

(*Manuscript accepted* 20 *May* 1967)

Br. J. Psychol. (1967), 58, 3 *and* 4, *pp.* 477–485
Printed in Great Britain

PUBLICATIONS REVIEWED

Textbook of Psychology. By D. O. Hebb. (2nd Ed.) London: Saunders. 1966. Pp. xi–353. 40*s*.

Textbook of Elementary Psychology. By E. Galanter. San Francisco: Holden-Day. 1966. Pp. xiii + 419. $8.75.

Both of these books have introductions which promise the student an uncompromising dose of science; pure science, since they both make the point that practical hints on how to live are no part of psychology and suggest that applied psychology is not quite intellectually respectable. The perceptive student may note that any real resemblance between the books ends with these declarations of faith; there is some overlap in content but none in style or treatment. A determined wish to kill a perfectly legitimate human interest which the student may bring to the study of psychology is evidently the first wish of both authors. Dr Galanter's method is the more direct, he simply advises the student who is 'interested in people' to talk to his grandmother; Dr Hebb achieves his effect partly by a deliberately mystifying use of the English language and partly by drawing most of the emotive material from chimpanzee rather than human behaviour (which is cheating).

In fact both authors suffer the same problem—that their books are not textbooks of psychology but rather textbooks of theories or branches of psychology. Both books are written from well-defined coherent points of view, which have considerable interest in their own right, but neither of which is co-terminous with psychology as a discipline. The evident determination to limit what the student is permitted to call psychology, rather than give the books more modest titles, seems on the face of it rather high-handed.

Dr Galanter is not very definite about whether his book is introductory—it is not; his own feeling that the student probably needs college mathematics and a course on computer logic before he starts the book is at least partly justified by the contents. On the other hand his further suggestion that after a course using the book 'it would be valuable for the student to enroll in a laboratory course' is merely depressing. No student who has not done, or is not doing, a laboratory course should use this book, or he will be grossly misled by the smoothed data in graphical form which is frequently used to illustrate the text. The student's grandmother might be able to tell him that a graph portraying the behaviour of a 'hypothetical animal' (for example, fig. VII–5) should be treated with reserve, but he had much better find out for himself the hard way in the laboratory.

Leaving these reservations aside, this is in many ways a masterly book which will take the student much further into the world of contemporary theory and experiment in psychology than Dr Hebb's text; it explains things like the theory of sets, information theory, probability models, linguistic theory and the treatment of evaluation problems in a single co-ordinated exposition of choice and the principles of organization which mediate it. Neither the theory nor the material is easy but it is clear and logical. The second-year student who makes himself familiar with this book will benefit greatly; it will sharpen his thinking and extend his knowledge.

Dr Hebb's book will serve a different purpose; he manages to catch in print a quality which is traditionally associated with demonstration or laboratory work—he gets the reader very close to the 'feel' and excitement of experiment and into the frame of reference where experiment arises from felt necessity. So few books come anywhere near this that this is still a good book on which to start students—whatever its limitations in other directions. Most of the material is psycho-physiological, with a few oddments from other areas. The argument flows quickly; the student will enjoy the vigorous individualistic approach of someone whose intellectual excitement is worth sharing; that this book should be so readable and yet so clearly an experimenter's book is no mean achievement.

Perhaps if Dr Hebb had omitted the glossary one could leave it at that and dismiss the lurking disquiet which his use of language provokes, but the glossary contains definitions which give rise to something stronger than mere disquiet. Two of the more disastrous redefinitions of common words will illustrate what is offered as pure science: jealously is defined as 'anger, characteristic

of certain social situations', and love as 'arousal accompanied by mediating processes tending to produce closer or continued exposure to the excitant (ideas of maintained contact with the object or person loved)'. To demand of students that they accept such nonsense as scientific is an insult to their intelligence; I only hope that no one is tempted to assault their integrity by asking them to say it is 'true' in an objective test.

MONICA LAWLOR

The Psychology of Learning. By R. BORGER and A. E. M. SEABORNE. Harmondsworth: Penguin. 1966. Pp. 250. 5*s*.

This book is clearly thought out and well expressed in straightforward language, lightened with agreeable touches of dry humour. Theory and experiment are related throughout to everyday experiences and problems, and to wider social and biological backgrounds. In the overall development of a topic, the authors distinguish clearly between hypotheses, experimental data confirming or infirming these, and reasonable but as yet unverified extrapolations.

The breadth and complexity of the field of learning are such that a book of this length could hardly do justice to all its aspects. At least one chapter on social learning—for example, the learning of attitudes, social norms, and role behaviour—would have been welcome, as would a more penetrating treatment of conceptual learning. Perhaps the most notable omission to a developmental psychologist is the absence of reference to play as a learning activity. In a thoughtful last chapter on 'Learning and education', where the authors again show their awareness of the wider issues at stake, they ask: 'But how do we make someone into an active learner—i.e. someone who has the ability to organize his own work, select and find his own sources, and who has the desire to do so?' This is just what children at play are doing: the problem is not how to make them do it to begin with, but how to prevent this approach to learning from being, perhaps irretrievably, lost in the process of 'education'.

Within the limits imposed by length, however, the authors treat a satisfactory range of topics, and treat them well. As a starting-point for further learning about learning, students should certainly buy it.

RICHARD R. SKEMP

Behaviour Therapy Techniques: A Guide to the Treatment of Neuroses. By JOSEPH WOLPE and ARNOLD A. LAZARUS. London: Pergamon. 1966. Pp. ix + 198. 21*s*.

The publication of this small book on behaviour therapy marks a point in time at which acceptance of this form of treatment permits presentation of certain main techniques to the intending practitioner. The aim has been to provide a guide 'to facilitate the process of education' in some important behaviour-therapy procedures, rather than to put forward a closely reasoned justification for their use. In short, this is a book primarily concerned with an outline of therapeutic practice, although the hope is expressed that the text can be supplemented by more formal training.

Much of the value of the book is in the clarity of presentation of steps and strategy in therapy, and the illustrative case-histories are both appropriate and convincing. This advantage is offset, as it must be in a text with the kind of limitations imposed by the authors, by a neglect of much important technical and theoretical detail, and this may lead many readers to assume an oversimplified and erroneous view of the complex problems involved.

In particular, very little attention is given to the theoretical issues involved—an omission which might well be acceptable were it not that the text often appears to imply that no serious difficulties exist at this level. Again, while there is some discussion of the results of behaviour therapy, the account given is not well balanced in that the evidence for the efficacy of this form of treatment is given undue weight.

On the other hand the book provides a number of useful, practical rules of thumb concerning the application of behaviour therapy which could only be acquired from long experience of the kind which Wolpe and Lazarus undoubtedly possess. In this special sense, i.e. by making available a kind of clinical 'know-how' in addition to a detailed description of therapeutic principles, this book makes a useful contribution to the existing literature.

H. R. BEECH

The Senses. By O. LOWENSTEIN. Harmondsworth: Penguin. 1966. Pp. 214. 5*s*.

Experiments in Visual Perception. Edited by M. D. VERNON. Harmondsworth: Penguin. 1966. Pp. 443. 8*s* 6*d*.

Prof. Lowenstein's book describes the mechanisms of sensory reception in man and other animals. The emphasis is on the comparative and quantitative aspects of the sense organs, and he gives a clear description of the nature of physical stimuli and their relation to events in the nervous system. The main classification of the senses is by the nature of the stimulus—electromagnetic, mechanical (including hearing) and chemical—and he makes good use of analogies with engineering mechanisms, including a chapter on automatic control mechanisms applied particularly to the muscle spindles and the labyrinth. The book is illustrated with clear figures although the four plates contribute little of value. For students, this is an excellent introductory book, covering a large amount of detail and theory in a very readable way, and providing the basic information necessary before going on to a study of the more psychological aspects of the senses and of perception.

Experiments in Visual Perception is a collection of shortened papers on human visual perception, almost all employing a psychological rather than a physiological approach to the subject. This is sufficient in some fields of perception such as the influence of social factors, attitude and motivation, but in other fields such as the perception of form, movement and colour the most interesting developments are those resulting from physiological studies, both at the receptor level and the recording from nerve cells at higher levels of the visual process. However, these developments and their influence on psychological experimentation are largely omitted from this book, even though some of these latter fields of perceptual research are covered. Within these limitations the papers included provide a reasonably good coverage of the field, although there are omissions, such as the absence of papers on geometrical illusions. The papers are classified by six main headings, followed by some extracts from Piaget's observations on his own children. The main headings are: form perception, including Koffka's introduction to Gestalt theory; space and distance, including experiments on various distortions of the visual field, and Julesz on binocular vision; constancy, movement perception, including Michotte and others on causality, and a comparison between J. F. Brown's Gestalt approach and the more physiological approach of Shipley *et al.*; set and attention, including some of Bruner's work, and also studies on sensory isolation and stabilized retinal images; motivation and personality, including studies on perceptual defence, and Tajfel on size and value. This book is mainly suitable for students who already have a suitable theoretical framework in which they can place these experimental studies.

ANDREW GREGORY

Grouping in Education. Edited by A. YATES. London: Wiley. 1966. Pp. 314. 96*s*.

The material for this book was provided by an international working party organized by the UNESCO Institute, Hamburg, consisting of individuals from England, France, Belgium, Sweden, Israel, Italy, Germany and the United States. 'Grouping' covers both inter-school and intra-school organization, and embraces selective systems of secondary education and also streaming and setting. The book is therefore particularly timely for this country at least. The preface sees the book as being 'stimulating to educational policy makers, school administrators and students of education from developing countries as well as developed countries, despite the fact that the material collected has been from developed countries'.

The contents fall into three parts. The first part is a survey of existing grouping practices, covering both allocation to schools and grouping within schools. This is a very useful account of national differences, and many British students will be grateful for the information on grading, tracking, setting and streaming. It is inevitable that some educational practices in the countries not represented in the working party receive less than adequate coverage: for example, clubs and societies get under a page, which will seem less than justice to educationists in the communist countries. And a very odd omission here—to British eyes—is that of 'houses'. The house system, borrowed from the public schools, is often a most significant feature of our maintained secondary schools, and one that has not always escaped criticism.

The third part of the book is a survey of research, in the form of abstracts of significant research from all the represented countries. This, to many, will be the most useful part of the book, and the reviewer, for one, is grateful for the summary of a number of researches that hitherto had escaped his notice. But how much more useful this section would have been had the abstracts been critical rather than merely descriptive! For example, the abstract of Douglas's 1964 research bears no hint that his results on streaming are vitiated by the effects of regression to the mean; and the reader of the abstract of Goldberg and Passow's otherwise excellent New York study in 1965 remains unaware of the limitations of their sample in terms of range of ability (a split at I.Q. 100 puts 2000 pupils above and only 200 below).

The middle section of the book considers factors affecting grouping practices, and the relation between such practices and educational and social goals. Here the book tends to fall below expectation. One would hope that discussion would be buttressed by reference to apposite research results. In considering the *cognitive* outcomes of selective schools and of streaming within the schools this, on the whole, does happen. The general outcome of research here is inconclusive, and the point is made that it is not the grouping itself that is important, but the use which is made of it. But it is in the consideration of *social* factors that this middle section leaves most to be desired. The familiar point is made that 'the school can minimize or maximize the cleavages within the larger society', and the author then argues that streaming not only maximizes such cleavages but is designed to do so. Now the amount of research which has been done on this aspect of grouping is far from large, but this is no reason for ignoring it. The only research report cited is one in support of the author's views, that of Willig (1963) in England. This involved 200 primary school pupils in six classes. The results, 'considered as a whole', suggest that there are 'social advantages of mixed ability grouping as opposed to streaming by academic attainment'. But Borg's 1964 Utah research, involving 2500 pupils aged 10–14, found that 'ability grouping... provides average and slow pupils a better chance of social recognition'; Wilcox (New York, 1961), investigating sixteen schools, found 'significant evidence of a more positive self-concept among pupils with I.Q.s below 90 when they were homogeneously grouped'; Taylor (California, 1952), with 1177 children from 38 schools, found 'a significant positive relationship between mental age and social acceptance'; Goldberg and Passow (New York, 1965), investigating 3000 children from 45 schools, found that 'for the slower pupils, broad range classes produced somewhat lower self-concepts and considerably lower "ideal" concepts than did the narrow range'.

The reviewer has cited these researches not in order to suggest that streaming *per se* produces more socially cohesive schools, but to underline the fact that there is at least *some* evidence which points in the opposite direction to Yates's thesis. And for the slow learner in particular. Just as with the cognitive outcomes, it is more sensible to believe that organization as such is relatively unimportant: what matters is the attitude of the school and its teachers to education and to the children in their charge. The belief of many (educationists as well as politicians) that the institution of comprehensive secondary schools will automatically produce dramatic effects in achievement and in social cohesiveness is difficult to understand. Nor is it any easier to agree with those who see the abolition of streaming and setting ushering in the social millennium. Few things in education are simple; outcomes are always 'multi-factorial'. It is unrealistic ever to expect a final and definitive research on 'grouping in education': the most we can do is to take a sober look at all the results available and endeavour to make a dispassionate judgement. This book falls short in this respect—but at least it provides much of the data for the reader to do it himself.

STEPHEN WISEMAN

Studies in Cognitive Growth. By J. S. Bruner, R. R. Olver and P. M. Greenfield, *et al*. New York: Wiley. 1966. Pp. xv–343. 60*s*.

This book gives an account of the work of Prof. Bruner and his colleagues into the processes of cognitive growth in children. The first two chapters, written by Bruner himself and occupying about one-fifth of the book, develop the views he outlined in a paper in the *American Psychologist* in 1964. These chapters argue that there are three ways in which the growing child can represent reality to himself: by action, by imagery and by symbolism in language. Very important, however, is the interaction between these modes of representation, for in Bruner's view, when these systems of representation are in conflict, cognitive growth is likely to occur. Hence the import-

ance of developing correspondence between what the child does, sees, and says. At the same time cognitive growth will be affected by the images, skills, concepts, etc., available in the culture, the way of life of the individual and the demands of the culture on him, and the extent to which he is motivated to explore the discrepancies among his three ways of knowing.

Chapters 3–13 report studies dealing with four broad issues involving growth: equivalence; the ways in which children seek information in efficient ways; conservation; equivalence and conservation under different cultural conditions. Individuals used in the studies included middle-class children in Massachusetts, Eskimo pupils in Alaska, rural and urban children in Mexico, and schooled and unschooled children in Senegal. While simple Whorfian notions about the relation between language and reality are rejected, stress is placed on the role of linguistic variables in conceptional growth.

The Harvard and Geneva schools are frequently in disagreement with one another. For example, Harvard believes that it is *identity* and not *reversibility* or *compensation* that turns actions into operations in the conservation problem; also that Piaget tends to ignore the place of imagery, and underestimates the role of language, in cognitive growth. A reply to Bruner *et al.* in respect of the relation between language and thought has already been made by Inhelder *et al.* in a letter in the *American Psychologist*, February 1966.

This is an important book which cannot be ignored by anyone concerned with cognitive development. It contains a great deal of experimental evidence, a wealth of ideas, and an approach to cognitive development which is complementary to that of the Geneva school. The volume must be studied with care. Readers must decide if modes of representation provide a more useful conceptual framework inside which to consider cognitive growth than does Piaget's framework of 'stages'. The authors write easily and readers must not allow the smoothness of the text to conceal possible weaknesses. For example, one does not expect in 1966 to read: 'At the Institute in Geneva, cognitive development is seen as almost purely a matter of maturation...' (p. 214)—this is far from the Geneva viewpoint. But more important to note—as so much of the subsequent argument depends on it—is the assumption (p. 229) that the schooled and unschooled groups of the Wolof tribe were 'equivalent with respect to native endowment'. Even if this is true we cannot be sure that parental attitudes were the same for the two groups. Studies of nursery and non-nursery children in this country are likewise bedevilled by the fact that although one equates for I.Q. and socio-economic status, background conditions are never sufficiently closely studied for one to be sure that it is the school which is the crucial variable, and that there are no differences between parents who send their children to school and those who do not. This is not to say that the authors are incorrect in respect of the value they place upon schooling in cognitive growth, but to point out that all other possible influences have not necessarily been excluded. In summary, this is an important contribution to the study of cognitive development, but the text needs close and constant scrutiny.

K. LOVELL

Anxiety and Behaviour. Edited by CHARLES D. SPIELBERGER. New York and London: Academic Press. 1966. Pp. xiv + 414. $9.75.

This book is really a collection of papers. With two exceptions, however, they are not reprinted but were specially written. A reader familiar with the field, skimming through the following paragraphs to pick out the names of contributors, will be able to predict the contents of most of the chapters fairly accurately.

In his preface the editor comments that 'the wide gap that separates contemporary personality theory from the empirical foundation upon which it must ultimately rest is painfully evident to the serious student of personality phenomena'. He goes on to say that he 'seeks to facilitate a more comprehensive understanding of anxiety and its effects on behaviour by bringing together the current views of psychologists and psychiatrists whose previous work and thought on this subject have contributed significantly to our present knowledge'. By 'bringing together' he apparently means 'assembling between the same covers' because there is no attempt at a synthesis. Spielberger himself admits that 'a meaningful synthesis of these diverse views is not yet possible'.

The book as a whole contributes little to closing the gap between theory and observation but some chapters, taken individually, are far better in this respect. One outstanding example is a chapter by Malmo giving a clear exposition of activation theory and the facts on which it is

based, with a useful account of the author's work comparing anxiety patients, early schizophrenics, and control subjects. There is little new evidence here, but Malmo does give us a valuable and coherent review of work published over many years as well as some fresh and stimulating thinking. The same can be said of the chapter by K. Spence and Janet Taylor Spence in which they set down with admirable clarity the latest work on their drive theory of anxiety in relation to performance. In another well-written chapter R. B. Cattell presents a useful discussion of the several anxiety factors that his work has revealed. A good chapter by Spielberger attempts with some success to apply Spence's drive theory to real-life achievement. Seymour B. Sarason raises some important questions concerning what is actually being measured by anxiety scales and there is an interesting contribution by Mandler and Watson on the anxiety-provoking effect of interrupting well-established behaviour patterns.

Schachter's chapter, reprinted from another book, presents interesting results on the effects of cognitive factors upon reported and observed emotions. It is unfortunately marred by the fact that the experiments reported involved a serious and deliberate deception of the students who took part concerning the nature of injections that they received. It is the reviewer's belief that these experiments should not have been done and that they are likely to be harmful to the long-term interests of all research workers who investigate human subjects.

Other chapters by R. R. Grinker, by R. S. Lazarus and E. M. Opton, and by C. W. Eriksen all contain useful and interesting material. Wolpe has a contribution which is a brief exposition of orthodox Wolpian theory and practice.

On the whole this is a useful publication, justified by the level of some of its contributions. They gain nothing from being published in a single volume, however, except the convenience of having them together in one place.

J. G. INGHAM

The Syndrome of Stuttering. By GAVIN ANDREWS and MARY HARRIS. London: Spastics Society and Heineman. 1964. Pp. 191. 40*s*.

This volume begins with a useful and fairly broad review of the literature on stuttering both in terms of aetiology and treatment. It accomplishes this end by sometimes covering the same work from different points of view. In association with this we are given details of a survey of stuttering in school-children plus appropriate control data—with particular reference to families, mothers and the development of the children as it seems to relate to the onset of stuttering. Finally some suggestive research is reported in which an attempt was made to inhibit stuttering by the use of syllable-timed speech. It is clear there is much in this volume which will be helpful, if not necessary, to the worker in this field. What astonishes this reviewer is that so little mention is made here of possible relations between stuttering and the general phenomena of language behaviour. Questions as to the possible effects of linguistic complexity—either syntactic or semantic—remain unasked. The nearest we get is a bald assumption that stuttering is only related to language in terms of the physical parameters of speech and hearing, and not to latent structure. The detailed analysis of interrelations between structural constraints on phonological regularities and the phenomena of stuttering might prove revealing both to workers in this field and to psycholinguistics in general. It might also shed some light on another assumption made by the authors, namely that unlearning of a particular behaviour pattern (e.g. stuttering) indicates that it was learned in the first place. A learning theory of reinforcement might be one which either accounts for reinforcement or one which uses reinforcement to account for something which is learned. These two theories raise different empirical questions—even about the stutterer.

ROGER J. WALES

New Horizons in Psychology. Edited by BRIAN M. FOSS. Harmondsworth: Penguin. 1966; reprinted 1967. Pp. 448. 7*s*. 6*d*.

The reprinted edition of this book is an improvement on the original. This included such gems in its table of contents as 'University of Newcastle upon Tyne, London' (carrying universities to London in return for coal?). The body of the book was better.

In many respects this Pelican is one of the best of the brood, and its twenty-one chapters constitute, as the editor hoped, a lively and stimulating 'progress report' in the more experi-

mental areas of psychology. 'It is intended also to be a guide to what may be some of the exciting developments in the coming years.' It should certainly serve to widen the horizons of students of psychology, the general reader, and even many teachers of psychology. The surveys, though variable in style, have all been well done and contain valuable information on recent developments in perception, thinking, communication, behaviour genetics, the physiology of motivation, sleep, learning, personality theory, group behaviour, etc.

Undoubtedly in these new horizons many readers will feel lost. Has the course of psychological research any direction? If we are not to be left in impotent amazement, sooner or later we shall need order and a sense of direction. Dr Bannister (p. 379) perhaps most clearly realizes this. 'It is difficult', he writes, 'to see how findings from different experiments are ever to be related or added to one another in any way unless they are derived from some common theoretical framework. The present position is that we have thousands of studies using different concepts and different operational definitions which represent so many mountainous piles of data which cannot be cross-referred. In the more advanced sciences it is generally agreed that the most useful way of achieving and expressing an understanding of an area is to account for events by principles which are embodied into an overall theory of wide applicability. There seems no reason why psychology as a science should be an exception.' Perhaps an overall theory may be beyond our present horizons. But our excitement at new facts should not dim our awareness of the need for theoretical underpinning. The union of theory and practice is not a bad principle to keep in mind.

L. S. HEARNSHAW

The Psychology of Union–Management Relations. By Ross Stagner and Hjalmar Rosen. London: Tavistock. 1965. Pp. vii + 147. 21*s*.

This particular example of the 'Behavioural Science in Industry' series represents an analysis of some psychological processes in industrial disputes and is designed primarily for advanced undergraduates and graduate students. On the first point the book succeeds, though I fear students of industrial relations are much more likely to find it enlightening than will psychologists.

The book is well written and illustrated, easy to read and covers the important psychological literature adequately. The authors work their way through some of the relevant fields of both psychology (perception, motivation, frustration and aggression) and industrial psychology (company and union as organizations, leadership). The usefulness of psychology in analysing or attempting to settle disputes in the field of perception, for example, may be relatively obvious but is nevertheless worth stressing. The problem in any attempt to resolve a dispute, therefore, is not to uncover 'the real facts' but 'to try to get each side to take a look at the facts as seen by the opposition'. There is a lesson here for the members of Royal Commissions entering this field as well as for union and management leaders. Use is made, on several occasions, of Lewin's 'field model', particularly to illustrate the origins of disagreement as we find, for example, in the union's primary concern for security of jobs and the management's interest in efficiency. Some psychologists are likely to quibble with the author's homeostatic theory of movitation, but in view of the widely recognized preoccupation of the unions with the deficiencies of their membership (pay, working hours, holidays, working conditions, etc.) this model is probably quite in order.

The two chapters on union–management relations in process are disappointing. Just when we are hoping for a cogent analysis of a number of studies or a few good case-studies presented in lively fashion, the authors confront us with brief and fairly trite statements on 'forms of disputes' and 'settling disputes'.

We are left with the feeling that while the authors' knowledge of psychology is considerable, their familiarity with the practical aspects of industrial relations is extremely limited.

DENIS PYM

Elementary Statistical Methods in Psychology and Education. By P. BLOMMERS and E. F. LINDQUIST. London: University of London Press. 1965. Pp. 528. 50*s*. Study Manual: Pp. 247. 17*s*. 6*d*.

Perhaps the first important point to be made is that the authors' avowed intent differs somewhat from that of most textbook writers. In their own words: 'It (i.e. the book) is not intended as a general reference book...instead, a relatively small number of basic statistical techniques and concepts have been developed much more thoroughly and systematically than is customary in texts with a wider topical coverage.' I suspect that this policy will tend to restrict the market for this book, since the authors are led to omit almost *in toto* all non-parametric statistics, together with analysis of variance. Because of the exclusion of so many very important and powerful statistical tests, the book could hardly stand alone as the basis for an introductory course in statistics.

On the other hand, coverage of various topics is very extensive and adequate. The conceptual background to the *t*-test, for example, is very clearly sketched in, and would almost certainly lead to a more enduring 'understanding' than that afforded by most textbooks in statistics.

Chapter 10 ('Testing statistical hypotheses') is another example of the excellence with which certain aspects of statistics are dealt with. This is the best account of hypothesis-testing I have come across, and, in the cliché, could be read with advantage by every student of statistics. Two pedagogical examples are selected. The first concerns the practical problem of whether the children of a school are 'on the whole' subnormal or not; the second is concerned with three experiments in learning theory, very clearly showing how the feedback obtained from the statistical analysis of one experiment helps to determine the next experiment. One can really see here how statistics meshes in with experimentation.

In fine, the emphases of the book seem slightly strange, and much important material is omitted; against this some parametric tests are afforded really full and coherent coverage. The study manual accompanying the book embodies a large range of problems, and appears fairly well integrated with the text.

M. W. EYSENCK

Age of Transfer to Secondary Education. By J. D. NISBET and N. S. ENTWISTLE. London: University of London Press. 1966. Pp. 111. 12*s*. 6*d*.

This is partly a report on the first stage of a research project, and partly a contribution to a current controversy within the larger debate on the organization of secondary education. The research involved is a follow-up of a group of children from the primary school through the age of transfer into secondary school. This is not yet completed, but results from it—where relevant—have been incorporated into this discussion.

This is a timely publication, coming as it does when the majority of local authorities in England and Wales are planning the change-over from the selective system of secondary education to some form of comprehensive schooling, and the Plowden Committee has put its recommendations into final shape. There is everything to be said for making the age of transfer flexible—at least from the administrative and organizational point of view. But what about its educational and psychological implications? What is the 'right' age of transfer?

These are the questions that Prof. Nisbet and Mr Entwistle attack in this book. Although of slender size—less than 100 pages excluding appendices—it gives us an admirable survey of the relevant considerations. It starts with a consideration of physical growth and the 'adolescent spurt', and their relationship to mental development. The available research is admirably and succinctly reviewed, and the authors conclude that (*a*) individual growth curves of intelligence show no evidence of a growth spurt at puberty; (*b*) sex differences in I.Q. are dependent on the test used; (*c*) the mental development of boys and girls appears to be parallel; (*d*) early maturers show a more or less consistent superiority from age 7 to age 15; the weight of evidence is against the existence of a mental spurt at puberty.

The next question examined is whether the age of transfer is, or ought to be, related to the onset of adolescence and its 'special needs'. The writings of Rosseau, Stanley Hall and even the Hadow Report are in this vein, but the authors here suggest that 'the argument that the approach of adolescence requires "a fresh departure in educational methods and organizations" must be

judged in the light of the drill and repetitive methods which were formerly assumed to be appropriate to the education of childhood'. Research suggests that the changes in emotional maturity and in motivation are gradual rather than sudden, and associated with chronological age more than with physiological maturity.

Two chapters with an historical slant examine, first, the reasons for the differences between the English and Scottish systems, and secondly some of the variations in European practice. It is a commonplace that a society creates the educational system that it wants, in order to sustain its own structure and function. In any country, the shape of educational organization is the product, deliberately or almost accidentally, of that country's previous political philosophy: historical forces are very strong. This is well brought out in the comparison between Scotland and England, and the discussion of the kind of considerations which led to a transfer age of 12+ in one country and 11+ in the other.

What, then, is the right age of transfer? The authors conclude by saying 'there is no one "correct age"...the transition from primary to secondary education should extend over the whole period from age 10 to age 13. These years should be regarded as a transitional period, during which there is a gradual change in curriculum and style of teaching.'

STEPHEN WISEMAN

The British Journal of Psychology

The British Journal of Psychology (Monograph Supplements)

The British Journal of Medical Psychology

The British Journal of Social and Clinical Psychology

The British Journal of Mathematical and Statistical Psychology

The British Journal of Educational Psychology (Jointly with A.T.C.D.E.)

The British Journal of Psychology (about 512 crown quarto pages per volume), *The British Journal of Medical Psychology* (about 350 crown quarto pages per volume), and the *British Journal of Social and Clinical Psychology* (about 320 crown quarto pages per volume) appear in quarterly parts, the subscription price of each, payable in advance, being £4. 10*s*. per volume post-free ($16.50 in the U.S.A.). Subscriptions to any of these journals may be sent to any bookseller, or subscription agent, or to the Cambridge University Press, Bentley House, P.O. Box 92, 200 Euston Road, London, N.W. 1. American enquiries should be addressed to the Cambridge University Press, American Branch, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 288 pages. The subscription rate, payable in advance, is £3. 10*s*. per volume post-free ($13 in the U.S.A.). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C. 4; or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number. The subscription price in the U.S.A. is $4.20. Subscriptions should be sent to the Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36 Essex Street, Strand, London, W.C. 2.

A series of *Monograph Supplements* is issued in connexion with *The British Journal of Psychology* but is not included in the subscription for that journal. Orders may be sent to any bookseller, or the Cambridge University Press, P.O. Box 92, 200 Euston Road, London, N.W. 1, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to Professor W. Sluckin, Department of Psychology, The University, Leicester;

for *The British Journal of Medical Psychology* should be sent to

Dr T. Freeman, Dundee Royal Liff Hospital, By Dundee;

for *The British Journal of Social and Clinical Psychology* should be sent to

Miss Thelma Veness (Social Psychology Editor), Department of Psychology, Birkbeck College, Malet Street, London, W.C. 1, *or* to
Dr R. R. Hetherington (Clinical Psychology Editor), 6 Abercromby Square, Liverpool 7;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor R. J. Audley, Department of Psychology, University College London, Gower Street, London, W.C. 1;

for *The British Journal of Educational Psychology* should be sent to

Professor J. D. Nisbet, Department of Education, University of Aberdeen, Old Aberdeen;

for *The Monograph Supplements* should be sent to

Professor I. M. L. Hunter, Department of Psychology, The University, Keele, Staffs., whose decision regarding the publication of all Monograph Supplements will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1 September of the preceding year.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £6. They are entitled to purchase any or all of the following Journals at special rates: *The British Journal of Psychology* (£1 annually), *The British Journal of Medical Psychology* (£1 annually), *The British Journal of Social and Clinical Psychology* (£1 annually), *The British Journal of Mathematical and Statistical Psychology* (£1. 10*s*. annually), and *The British Journal of Educational Psychology* (£1 annually).

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, Tavistock House South, Tavistock Square, London, W.C. 1.

THE BRITISH JOURNAL OF

PSYCHOLOGY

EDITED BY

ARTHUR SUMMERFIELD

ASSISTANT EDITORS

JAMES HOGG PETER MITTLER

VOLUME 58, 1967

CAMBRIDGE
AT THE UNIVERSITY PRESS
1967

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York, N.Y. 10022

Printed in Great Britain at the University Printing House, Cambridge

CONTENTS OF VOLUME 58

Parts 1 and 2. February 1967

PAGE

An information-processing explanation of some perceptual phenomena. By H. A. Simon 1

Sensory spatial after-effects in relation to chronological age, mental retardation and schizophrenia. By R. H. Day, R. Burns, G. Singer, V. Holmes and D. Letcher 13

Adapting Witkin's field independence theory to accommodate new information from Africa. By M. Wober 29

Some experiments relating to Bartlett's theory of remembering. By A. Gauld and G. M. Stephenson 39

Affective connotation of words as a factor in free recall learning. By J. A. McNulty and C. J. E. Isnor 51

Digram-frequency and immediate memory. By M. J. A. Howe 57

Repetition and immediate memory. By E. C. Dalrymple-Alford . . . 63

Information processing related to stimulus novelty and complexity in a signal detection paradigm. By B. Weiner and P. Feldman 69

Effects of arousing stimulation upon novelty preference in rats. By H. C. Haywood and T. D. Wachs 77

Deutsch's model applied to a conditional maze. By D. H. Holding and D. Pritchatt 85

Basal conductance level and motor performance. By R. S. Corteen . . . 93

Sensory processing and recognition of forms in nursery-school children. By B. Denner and S. Cashdan 101

Conditioned orienting reactions in two-month-old infants. By J. Koch . 105

Rigidity, drive and conditioning in neurotics. By R. W. Payne, C. T. Netley and R. B. Sloane 111

The interaction of noise and personality with critical flicker fusion performance. By C. D. Frith 127

Musical taste and personality. By E. Payne 133

Individual differences in pain threshold and level of arousal. By D. R. Haslam 139

Note on the distribution of dominance times in binocular rivalry. By W. J. M. Levelt 143

A note on intelligence and season of birth. By G. W. Berglund 147

The genetic determination of intelligence: a reply. By C. Burt 153

PAGE

Indiscriminate operationism: a reply to Keehn on 'Avoidance responses as discriminated operants'. By R. Brown 163

Publications reviewed 167

Parts 3 and 4. November 1967

Serial and parallel classification. By D. W. J. Corcoran 197

Meaning mediation in verbal transfer. By R. Kanungo 205

Remembering of words by psychotic and subnormal children. By B. Hermelin and N. O'Connor 213

Memory after reading aloud and reading silently. By E. C. Poulton and C. H. Brown 219

Structure and content in concept attainment: effects of varying problem content. By N. E. Wetherick 223

A frequency analysis of controlled, continuous word association. By W. A. Matthews 227

Phrase units and the recall of grammatically structured nonsense. By P. Herriot 237

Order of retrieval of information simultaneously presented to ear and finger. By C. Madsen and L. Sims 243

Visual imagery and the organizational properties of perceived stimuli. By P. W. Sheehan 247

An analogue of extraversion as a determinant of individual differences in behaviour in the rat. By E. Weldon 253

Some problems in repertory grid measurement: I. The use of bipolar constructs. By J. M. M. Mair 261

Some problems in repertory grid measurement: II. The use of whole-figure constructs. By J. M. M. Mair 271

The relation of verbal rate to syllable length of verbal action in two interview situations. By J. T. Webb and D. M. Brister 283

Age differences in the temporal integration of language. By F. I. M. Craik and P. A. Masani 291

Water, fog and the size–distance invariance hypothesis. By H. E. Ross . 301

Further studies of pattern perception and a stabilized retinal image: the use of prolonged after-images to achieve perfect stabilization. By C. R. Evans 315

Adaptation to prisms: change in internally registered eye-position. By B. Craske 329

Stereoscopic vision and aniseikonic lenses. I. By G. N. Seagrim . . . 337

PAGE

Stereoscopic vision and aniseikonic lenses. II. By G. N. SEAGRIM . . . 351

Effects of a novel situation and of anxiety on two groups of dependency behaviours. By M. K. ROSENTHAL 357

Temporal effects on response decrement and stimulus satiation in exploration. By C. HUTT 365

Components of curiosity: Berlyne reconsidered. By E. GREENBERGER, J. WOLDMAN and S. W. YOURSHAW 375

Effect of delay of escape reward and response compatibility on human reaction time. By A. KARRAS 387

An examination of test situations involving suggestibility by indirection. By G. SINGER 395

A comparison of visual and pseudo-recognition thresholds as a function of frequency and recency. By S. PYKE, N. McK. AGNEW and V. PEZER . 403

Experimental study of the effect of letter reversals on reading. By D. ASSO and M. WYKE 413

Recent developments in short-term memory. By B. B. MURDOCK, JR. . . 421

The role of language and experience on the use of logical symbols. By J. YOUNISS and H. G. FURTH 435

The contribution of epiphenomenalism: a response to E. A. Locke. By W. A. NORTON 445

The nature of epiphenomenalism: a reply to E. A. Locke. By M. J. A. GLICKMAN 449

Epiphenomenalism and the contradictions of its critics. By H. J. EYSENCK . 451

A note on trends in answers to Cattell personality questionnaires by Scottish subjects. By J. MCQUAID 455

A note on field dependency and movement after-effects. By E. BAKER and N. F. DIXON 459

Simple apparatus for multidimensional vibrotactile signalling. By D. D. DIESPECKER 463

'Brain and conscious experience': a critical notice from the U.S.S.R. of the symposium edited by J. C. Eccles. By A. R. LURIA 467

Publications reviewed 477

INDEX OF AUTHORS

	PAGE
Agnew, N. M.	403
Asso, D.	413
Baker, E.	459
Berglund, G.	147
Brister, D. M.	283
Brown, C. H.	219
Brown, R.	163
Burns, R.	13
Burt, C.	153
Cashdan, S.	101
Corcoran, D. W. J.	197
Corteen, R. S.	93
Craik, F. I. M.	291
Craske, B.	329
Dalrymple-Alford, E. C.	63
Day, R. H.	13
Denner, B.	101
Diespecker, D. D.	463
Dixon, N. F.	459
Evans, C. R.	315
Eysenck, H. J.	451
Feldman, P.	69
Frith, C. D.	127
Furth, H. G.	435
Gauld, A.	39
Glickman, M. J. A.	449
Greenberger, E.	375
Haslam, D. R.	139
Haywood, H. C.	77
Hermelin, B. M.	213
Herriot, P.	237
Holding, D. H.	85
Holmes, V.	13
Howe, M. J. A.	57
Hutt, C.	365
Isnor, C. J. E.	51
Kanungo, R.	205
Karras, A.	387
Koch, J.	105
Letcher, D.	13
Levelt, W. J. M.	143
Luria, A. R.	467
McNulty, J. A.	51
McQuaid, J.	455
Madsen, M. C.	243
Mair, J. M. M.	261, 271
Masani, P. A.	291
Matthews, W. A.	227
Murdock, B. B., Jr.	421
Netley, C. T.	111
Norton, W. A.	445
O'Connor, N.	213
Payne, E.	133
Payne, R. W.	111
Pezer, V.	403
Poulton, E. C.	219
Pritchatt, D.	85
Pyke, S.	403
Rosenthal, M. K.	357
Ross, H. E.	301
Seagrim, G. N.	337, 351
Sheehan, P. W.	247
Simon, H. A.	1
Sims, L.	243
Singer, G.	13, 395
Sloane, R. B.	111
Stephenson, G. M.	39
Wachs, T. D.	77
Webb, J. T.	283
Weiner, B.	69
Weldon, E.	253
Wetherick, N. E.	223
Wober, M.	29
Woldman, J.	375
Wyke, M.	413
Youniss, J.	435
Yourshaw, S. W.	375

INDEX OF PUBLICATIONS REVIEWED

ANDREWS, G. & HARRIS, M. *The Syndrome of Stuttering* 482
ANSTEY, E. *Psychological Tests* 168
BILODEAU, E. A. (ed.). *Acquisition of Skill* 180
BLOMMERS, P. & LINDQUIST, E. F. *Elementary Statistical Methods in Psychology and Education* 484
BORGER, R. & SEABORNE, A. E. M. *The Psychology of Learning* 478
BOURNE, L. E., JR. *Human Conceptual Behaviour* 167
BRUNER, J. S., *et al. Studies in Cognitive Growth* 480
CONGER, J. J. & MILLER, W. C. *Personality, Social Class and Delinquency* 169
COSTELLO, C. G. *Psychology for Psychiatrists* 186
DEUTSCH, M. & KRAUSS, R. *Theories in Social Psychology* 170
ELLIS, H. C. *The Transfer of Learning* 184
FELDMAN, S. (ed.). *Cognitive Consistency* 193
FOSS, B. M. (ed.). *New Horizons in Psychology* 482
GAITO, J. (ed.). *Macromolecules and Behaviour* 181
GALANTER, E. *Textbook of Elementary Psychology* 477
GIBBONS, T. C. N. & AHRENFELDT, R. H. *Cultural Factors in Delinquency* 176
GLASER, E. M. *The Physiological Basis of Habituation* 182
HEBB, D. O. *Textbook of Psychology* (2nd ed.) 477
HEBRON, M. E. *Motivated Learning: A Developmental Study from Birth to the Senium* 173
HORDEN, A., BURT, C. & HOLT, N. F. *Depressive States: A Pharmacotherapeutic Study* 176
HUDSON, L. *Contrary Imaginations: A Psychological Study of the English Schoolboy* 173
HURVICH, L. M. & JAMESON, D. *The Perception of Brightness and Darkness* 191
INGLIS, J. *The Scientific Study of Abnormal Behaviour* 193
LAING, R. D., *et al. Interpersonal Perception* 191
LORR, M. (ed.). *Explorations in Typing Psychotics* 179
LOWENSTEIN, O. *The Senses* 479
MCGRATH, J. E. & ALTMAN, I. *Small Group Research: A Synthesis and Critique of the Field* 170
MAGNUSSON, D. *Test Theory* 187
MAHER, B. A. *Principles of Psychopathology. An Experimental Approach* 171
MANDLER, G., *et al. New Directions in Psychology, III* 177
MARLER, P. R. & HAMILTON, W. J. *Mechanisms of Animal Behaviour* 189
MEADE, J. E. & PARKES, A. S. (eds.). *Genetic and Environmental Factors in Human Ability* 185
MILES, T. R. *Eliminating the Unconscious* 171
MORGENSTERN, M., *et al. Practical Training for the Severely Handicapped Child* 190
NISBET, J. D. & ENTWISTLE, N. S. *Age of Transfer to Secondary Education* 484
OPPENHEIM, A. N. *Questionnaire Design and Attitude Measurement* 192
PRINGLE, M. L. K. *Adoption—Facts and Fallacies* 188
RAO, K. R. *Experimental Parapsychology* 183
RAVEN, J. C. *Psychological Principles appropriate to Social and Clinical Problems* 169
ROZEBOOM, W. W. *Foundation of the Theory of Prediction* 187
DE RUECK, A. & KNIGHT, J. (eds.). *Conflict in Society* 182
RUSSELL, R. W. (ed.). *Frontiers on Physiological Psychology* 194
SEGALL, M. H., *et al. The Influence of Culture on Visual Perception* 188
SPIELBERGER, C. D. (ed.). *Anxiety and Behaviour* 481
STAGNER, R. & ROSEN, H. *The Psychology of Union–Management Relations* 483
STERNBACH, R. A. *Principles of Psychophysiology* 183
STOTT, D. H. *Studies of Troublesome Children* 175
UZNADZE, D. N. *The Psychology of Set* 167
VERNON, M. D. (ed.). *Experiments in Visual Perception* 479
WHITE, A. R. *Attention* 186
WOLPE, J. & LAZARUS, A. A. *Behaviour Therapy Techniques: A Guide to the Treatment of Neuroses* 478
WOODBURNE, L. S. *The Neural Basis of Behaviour* 192
YATES, A. (ed.). *Grouping in Education* 479
YINGER, J. M. *Toward a Field Theory of Behavior: Personality and Social Structure* 170

KARRAS, ATHAN
Effect of delay of escape reward and response compatibility on human reaction time 387

SINGER, G.
An examination of test situations involving suggestibility by indirection 395

PYKE, SANDRA, AGNEW, NEIL MCK. and PEZER, VERA
A comparison of visual and pseudo-recognition thresholds as a function of frequency and recency 403

ASSO, DOREEN and WYKE, MARIA
Experimental study of the effect of letter reversals on reading 413

MURDOCK, BENNET B., JR.
Recent developments in short-term memory 421

YOUNISS, JAMES and FURTH, HANS G.
The role of language and experience on the use of logical symbols 435

NORTON, W. A.
The contradiction of epiphenomenalism: a response to E. A. Locke 445

GLICKMAN, MAURICE J. A.
The nature of epiphenomenalism: a reply to E. A. Locke 449

EYSENCK, H. J.
Epiphenomenalism and the contradictions of its critics 451

MCQUAID, JOHN
A note on trends in answers to Cattell personality questionnaires by Scottish subjects 455

BAKER, E. and DIXON, N. F.
A note on field dependency and movement after-effects 459

DIESPECKER, D. D.
Simple apparatus for multidimensional vibrotactile signalling 463

LURIA, A. R.
'Brain and conscious experience': a critical notice from the U.S.S.R. of the symposium edited by J. C. Eccles (1966) 467

PUBLICATIONS REVIEWED 477

22 JUL 1968

The British Journal of Psychology *Volume 58, Parts 3 and 4, 1967*

U·A

CONTENTS

CORCORAN, D. W. J.
Serial and parallel classification *page* 197

KANUNGO, R.
Meaning mediation in verbal transfer 205

HERMELIN, BEATE and O'CONNOR, N.
Remembering of words by psychotic and subnormal children 213

POULTON, E. C. and BROWN, C. H.
Memory after reading aloud and reading silently 219

WETHERICK, N. E.
Structure and content in concept attainment: effects of varying problem content 223

MATTHEWS, W. A.
A frequency analysis of controlled, continuous word association 227

HERRIOT, P.
Phrase units and the recall of grammatically structured nonsense 237

MADSEN, MILLARD C. and SIMS, LANNY
Order of retrieval of information simultaneously presented to ear and finger 243

SHEEHAN, PETER W.
Visual imagery and the organizational properties of perceived stimuli 247

WELDON, E.
An analogue of extraversion as a determinant of individual differences in behaviour in the rat 253

MAIR, J. M. M.
Some problems in repertory grid measurement: I. The use of bipolar constructs 261

MAIR, J. M. M.
Some problems in repertory grid measurement: II. The use of whole-figure constructs 271

WEBB, JAMES T. and BRISTER, DAVID M.
The relation of verbal rate to syllable length of verbal action in two interview situations 283

CRAIK, F. I. M. and MASANI, P. A.
Age differences in the temporal integration of language 291

ROSS, HELEN E.
Water, fog and the size–distance invariance hypothesis 301

EVANS, C. R.
Further studies of pattern perception and a stabilized retinal image: the use of prolonged after-images to achieve perfect stabilization 315

CRASKE, B.
Adaptation to prisms: change in internally registered eye-position 329

SEAGRIM, G. N.
Stereoscopic vision and aniseikonic lenses. I 337

SEAGRIM, G. N.
Stereoscopic vision and aniseikonic lenses. II 351

ROSENTHAL, MIRIAM K.
Effects of a novel situation and of anxiety on two groups of dependency behaviours 357

HUTT, CORINNE
Temporal effects on response decrement and stimulus satiation in exploration 365

GREENBERGER, ELLEN, WOLDMAN, JANET and YOURSHAW, SARAH WILLIAMS
Components of curiosity: Berlyne reconsidered 375

Continued inside back cover

Printed in Great Britain at the University Printing House, Cambridge